DOCUMENT RESUME

ED 151 416 95 TM 006 946

AUTHOR Watkins, Richard W.; Ruyle, Janet
TITLE Information Specifications for Development of Instruments to be Used in the Evaluation of Nontraditional Programs.
INSTITUTION California Univ., Berkeley. Center for Research and Development in Higher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Jun 77
CONTRACT 400-76-0107
NOTE 118p.; For related document, see ED 112 712

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS College Programs; College Students; Cost Effectiveness; Data Analysis; *Data Collection; *Educational Assessment; *Evaluation Criteria; Evaluation Needs; External Degree Programs; *Information Needs; Post Secondary Education; *Program Evaluation; Research Design; *Special Degree Programs; Student Evaluation; Surveys; Test Construction

ABSTRACT

Specifications for the types of information necessary for the evaluation of nontraditional college programs involve: content, respondent groups, schedule for data collection, possible instruments and forms of response, and presentation of the information to different groups. One important issue is the need for more precise definitions of traditional and nontraditional degree programs. Three general categories of nontraditonal programs have been identified in order to suggest some of the problems encountered in the development of information collection procedures. These programs include: (1) on-campus curricula and degrees extended by flexible scheduling and location, (2) periodic short-term residence alternating with longer periods of self-disciplined study, and (3) individually developed degree programs defined by contract agreements between the student and the program faculty. This report is limited to the specifications for information about associate and baccalaureate degree programs. (Author/MV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151416

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

INFORMATION SPECIFICATIONS FOR DEVELOPMENT OF INSTRUMENTS TO BE USED IN THE EVALUATION OF NONTRADITIONAL PROGRAMS

Richard W. Watkins

Janet Ruyle

Center for Research and Development in Higher Education
University of California, Berkeley

June 1977

TM006 946

The research reported herein was supported by Contract No. 400-76-0107 from the U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position or policy of the National Institute of Education.

INFORMATION SPECIFICATIONS FOR DEVELOPMENT OF INSTRUMENTS TO BE USED IN THE EVALUATION OF NONTRADITIONAL PROGRAMS

Richard W. Watkins

Janet Ruyle

Center for Research and Development in Higher Education
University of California, Berkeley

June 1977

Foreword

TOWARD AN EVALUATION OF NONTRADITIONAL DEGREE PROGRAMS

In August 1976, the Center for Research and Development in Higher Education at Berkeley began work on a nine-month contract with the National Institute of Education to plan a two-year evaluation study of nontraditional degree programs. After five months of work during which the Center staff responded to a variety of propositions set forth by NIE in its original proposed statement of work, it became clear that many factors made it impossible to put an evaluation study in place by the May 15 deadline.

In January 1977, NIE and the staff agreed that for the remainder of the contract period it would be more productive for the staff to complete three miniprojects related to the evaluation of nontraditional degree programs rather than to plan a specific study as originally envisaged. The three miniprojects are:

I. The Needs of Nontraditional Learners and the Responses of Nontraditional Programs

II. Identification of Decisionmakers Concerned with Nontraditional Degree Programs and an Analysis of Their Information Needs

III. Information Specifications for Development of Instruments to be Used in the Evaluation of Nontraditional Programs

The miniprojects, while all related to the evaluation of nontraditional programs, were written as three separate reports. It is hoped that the miniprojects will provide useful background for reaching eventual agreements on some appropriate evaluation designs for the enormous diversity represented in "nontraditional education."

Contents

		Page
	Foreword: Toward an Evaluation of Nontraditional Degree Programs	i
I.	Miniproject Procedures and Information Priorities	1
II.	Study Design and Analysis Assumptions	7
III.	Information Specifications	16
IV.	Conclusions and Recommendations	86
	References	93
	Appendix	94

I. MINIPROJECT PROCEDURES AND INFORMATION PRIORITIES

The objective of this miniproject is to provide a first statement of specifications for information necessary to be responsive to the major needs of decisionmakers. The specifications generally deal with the following information characteristics: content, respondent groups, schedule for collection, possible instruments and forms of response, and presentation of infcrmation for different groups. Preparing information specifications is only an early stage in the preparation of a plan for data collection and analysis. The objective of this miniproject intentionally omitted data collection planning because there are so many important issues that cannot be resolved until the actual planning phase for a defined study.

Ideally, this miniproject should have followed the miniproject *Identification of Decisionmakers and an Analysis of Their Information Needs*. Practically, they had to be done concurrently. Although the two miniprojects were carried out by different staff members, information and experiences were shared during the course of the work. Additionally, one set of institutional visits was made jointly by staff members of each of the two projects.

One issue of overriding importance is the need for a more precise delineation of "traditional" and "nontraditional" degree programs. These terms have been used throughout this report as if they had meaning, because they were used in the original scope of work statement. In the second section of this report, three very general

categories of nontraditional programs have been identified in order to suggest some of the detailed problems that will be encountered in the development of information collection procedures. No similar categorization of traditional programs has been attempted, although there is great variety among traditional programs. Any definitions of traditional and nontraditional programs will be complicated and in the final analysis may have to be quite arbitrary in many respects. It did not seem necessary to make the arbitrary distinctions in order to consider the kinds of information needed for such an evaluation study.

Finally, this report is limited to specifications for information about students seeking associate and baccalaureate degrees as proposed in the original scope of work. The limitation shaped the program outcomes specifications in important ways, and also influenced the process information specifications. These specifications will be of limited usefulness for the evaluation of other kinds of educational activities or programs designed for adults.

PROCEDURES

A variety of activities was undertaken in the effort to produce specifications for information that could be both responsive to the identified needs of the largest number of decisionmakers and at least potentially feasible for an external agency to collect. These activities are listed below in approximately chronological order. Some of them overlapped, however, and others were carried out intermittently.

1. Review of evaluation reports of various kinds produced by a range of nontraditional programs. The major purpose of these reviews was to learn about the kinds of evaluation questions implied in the reports, the kinds of information obtained, and the forms in which the resulting data were presented.

2. Initial telephone contacts and some face-to-face interviews with faculty and staff members having responsibility for multiprogram evaluation of innovative programs at the systems level.

3. Interviews with staff of four nontraditional degree programs. The interviewees included college presidents, program administrators, institutional research specialists, and instructional staff. The focus was on the kinds of information obtained, the information collection devices, and the strengths and weaknesses of available information.

4. Attendance at two workshops designed for personnel with responsibility for evaluation of higher education programs. The first of these was the *Workshop on Evaluation of External Degree Programs* sponsored by California State University and Colleges Commission on External Degree Programs. This was one of two one-day workshops, and was attended by representatives from eight external degree programs in northern California. In addition to the formal presentations, the workshop provided an opportunity to talk with the program representatives and staff of the CSUC Chancellor's Office about information needs and collection problems.

The second workshop was titled *Outcomes Assessment Workshop* sponsored by the National Center for Higher Education Management Systems. Again, the workshop provided an opportunity to talk with institutional research staff from a number of two- and four-year colleges and universities. The materials also provided a number of references to data collection devices keyed to decisionmakers' information priorities.

5. Continuing assembly and review of evaluation priorities lists, models, and plans.

6. Consideration of crucial assumptions and constraints relevant to evaluation design and data analysis issues that would operate in the use of information specified.

7. Review and classification of items and questions in data collection forms used for evaluation of higher education programs--both traditional and nontraditional. These items were tentatively grouped and regrouped as the specifications were drafted and revised, to see where clarification, restatement, additional specifications, or reorganization were needed. The process was also helpful in increasing awareness of the extent of variation in the kinds of program features about which program and other staff are seeking information. The instruments used most often for this purpose are listed in the appendix.

QUESTION AND INFORMATION PRIORITIES

Decisionmakers functioning in different roles and at different levels of responsibility have different information needs, although there is a considerable amount of overlap in their needs. These differences have been documented and discussed in the TEAM and NCHEMS reports (Smith & Smith, 1975; Micek & Arney, 1974). While the TEAM report is based on information desired by decisionmakers about non-traditional education, the NCHEMS report ranks preferences for information on higher education outcomes in general.

It is not conceivable that a single evaluation study of even a relatively small number of programs--traditional or nontraditional--could satisfy well the major information needs of all those who must make decisions affecting the programs. There does appear to be some consensus on the areas of concern among those decisionmakers not directly involved in the administration and conduct of particular education programs, however; the differences among these decision-makers are in the priorities they assign to the areas and the kinds of information and amount of detail desired. It seemed best, therefore, to base the organization of the information specifications first on the consensus about major questions or areas of concern. The differences among decisionmaker groups could be reflected in the details of the information specifications and possible analyses.

The questions on which there is a consensus on high priority are presented below. For each question, a very general indication of the kinds of information that might be obtained is presented as an aid in defining the questions. It should be noted that this listing represents a judgmental pooling of a variety of indications of priorities rather than the results of specific studies.

1. Does the program satisfy learner needs for personal and professional development? Who are the dropouts? Why did they drop out? (Outcomes)

Types of information

- Satisfaction with knowledge and skills acquired

- Satisfaction with progress toward educational and career objectives
- Developed knowledge and skills
- Documented educational and occupational career development

2. Who are the interested and actual learners? Why did students enroll in the program? (Students and program procedures)

Types of information

- Demographic and personal characteristics
- Prior educational and occupational experience
- Measured skills and knowledge at entrance
- Educational and occupational objectives
- Expectations for learning methods, locations, and times

3. How much do the programs cost? Are they financially viable? (Costs)

Types of information

- Program costs
- Tuition, fees, and private costs to students
- Historical and projected sources of funds

4. How well do different program features address the needs of learners? (Program procedures)

Types of information

- Satisfaction with relevance of each feature to educational objectives
- Satisfaction with faculty, staff, materials, and resources of program
- Satisfaction with choices of time, place, and methods for learning

5. Are the degrees credible and the procedures for assessment of learning valid and reliable? (Faculty and administrators)

Types of information

- Acceptance of degrees and credits by other institutions
- Accreditation by external agencies or by groups within the institutions

- Student and faculty satisfaction with procedures for assessment of learning

6. To what extent does the program impact on other postsecondary programs? (Impact)

Types of information

- Adoptions or adaptations by other programs/institutions of materials, methods, or procedures developed by the program
- Modifications or adaptations of faculty procedures or courses to accommodate special student or program needs

The consensus on the ordering of these questions and information types is greater at the state and institutional/system administration levels of decisionmaking than at the program administration/faculty level. The latter group may well give highest priority to Questions 4 and 5. Much of the information they need for program decisionmaking requires a level of detail that can best be satisfied by analyses of data collected specifically by the staff of each individual program. Collection and analysis of this necessarily detailed information as part of a multiprogram evaluation study would produce an unmanageable amount of data.

II. STUDY DESIGN AND ANALYSIS ASSUMPTIONS

Information specifications cannot be derived entirely from a consideration of decisionmakers' information needs. There are too many options for programs and student groups for which information can be obtained. Therefore, some working assumptions had to be made about the major features of a data collection and analysis plan in preparing specifications for information. To arrive at these working assumptions, a model describing the interaction of students, program staff, procedures, and costs was chosen. Additionally, some reasonable constraints on the length of time over which the data collection effort might extend were set.

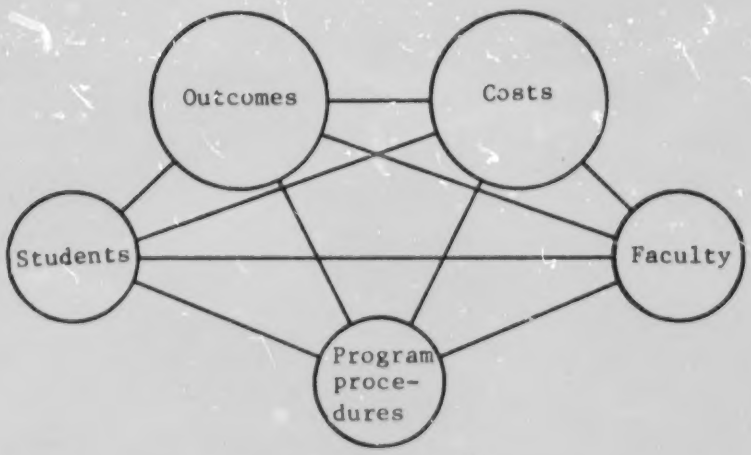
In considering the time constraints under which an evaluation study would have to be conducted, it seemed evident that any useful evaluation study would have to include the collection of longitudinal information for students, but that a study that followed both full- and parttime adult learners from start to completion of a baccalaureate degree would require too much time. A decision was made, therefore, that the longitudinal study of students would have to be done as a set of three essentially independent two-year longitudinal substudies. This time period was, in fact, incorporated in the original work statement and planning project proposal.

Since students are the focus of any educational program, student- or alumni-based outcomes information should be the highest priority information to be sought. The opportunities for learning occur in a context that includes many important factors and interests, however,

and the student outcomes result from the interaction of all these factors. The complexity of these interactions are described in Figure 1, a modified version of a figure first presented by Palola and Lehman (1977).

Figure 1

Student-Program Interaction Model



The high priority questions and relevant information listed in Section I can be related to this model, and analyzed in terms of the implications for the designation of groups from whom information should be sought.

Outcomes--Question 1

(a) for those progressing towards or receiving degrees
(b) for those who received degrees in prior years
(c) for those who showed interest in degree work but did not pursue it to completion

Students--Question 2

(a) characteristics when applying and/or entering
(b) educational and degree objectives

Costs--Question 3

(a) program costs and funding
(b) additional private costs to students

Program procedures--Questions 2 and 4

(a) description, classification, and assessment of procedures in response to student needs

Faculty--Question 5

(a) assessment of quality of procedures
(b) identification of strengths, weaknesses, and problems in procedures

This analysis suggested that information should be obtained from four groups of "students": those interested who never enrolled or who terminated their work prior to receipt of degrees, those seeking associate degrees, those seeking baccalaureate degrees, and those who received degrees in prior years (alumni). Additionally, information should be obtained from two groups of professional staff.

The student-program interaction model is most appropriate for study of a single program. Application of the model to a multiprogram study poses additional problems that must be considered before the general features of an evaluation study plan can be determined.

The fact that no two colleges or degree programs are identical cannot be disputed. Many apparent differences may be trivial, but many more are important or even crucial. The degree of importance of these differences and the details of them are most often not known or well understood, but they do exist.

This truism would not merit statement were it not for the fact that discussion of decisionmaking and of planning for multiprogram evaluation studies virtually require that the statement be overlooked. If an innovation that is even moderately different from most existing programs is to be evaluated in a multiprogram study, it is simply not possible to consider all the important variations among the several innovative and existing programs. If decisions affecting both established and alternative programs must be made, time and other pressures force a grouping of quite different programs into a very limited number of categories; for example, traditional and nontraditional, liberal arts and occupationally oriented, and two-year and four-year colleges.

The differences among students and student groups complicates the problem further. In many instances, different degree programs were designed for a variety of potential students. Nontraditional degree programs are the prime example of this, but not the only one; traditional degree-granting institutions have also introduced marked variations in program procedures to meet the special needs of various groups of students.

Any consideration of the most important interactions between students grouped on combinations of characteristics and programs grouped on combinations of major features quickly produces more subgroups for analysis than can be dealt with explicitly in an evaluation study. Additionally, variation on factors relevant to learning among students in each of the programs may be as important as the variation among programs designed for different kinds of students.

Any feasible multiprogram information collection plan will have to overlook many of the important differences among the students in these groups. The experimental design procedure of random assignment is clearly not appropriate; statistical control or analyses for separate subgroups can, at best, account for only a small part of these differences.

SUBSTUDY DEFINITION

These observations led to the conclusion that if a multiprogram evaluation is to be conducted, the several programs would have to be grouped in ways that would permit some pooling of data for similar programs, both traditional and nontraditional, and for some appropriate comparisons between dissimilar programs designed to accomplish reasonably common objectives. The phrase "reasonably common objectives" clearly implies comparison in terms of rather general long-term objectives; if the programs are dissimilar, it is due in part to the fact that they have different short-term objectives.

A number of schemes for grouping both traditional and nontraditional programs have been developed. The more sophisticated schemes

would probably require preliminary data collection about students attracted to various programs. The choice of a scheme appropriate to the final study plan will be an issue of major importance.

Recognizing that any grouping of different programs is a gross oversimplification of the complexities of each of the programs that may do violence to important features of them, three types of nontraditional programs were identified for purposes of determining the general requirements for information collection and analysis. The types are those described by Medsker et al. (1975).

1. On-campus curricula and degrees extended by flexible scheduling and location. Examples are the External Degree Program (California State University and Colleges), Armed Forces Contract Degree Program (Chapman College).

2. Periodic short-term residence alternating with longer periods of self-disciplined study, following a program-prescribed course of study. Examples are the Bachelor of Liberal Studies Program (University of Oklahoma), Bachelor of Arts in Liberal Studies (State University College of New York at Brockport).

3. Individually developed or negotiated degree programs of study defined by a contract or series of agreements between a student and the program. Examples are Empire State College, Metropolitan State College, Florida State University System External Degree Program (Florida International University).

The specifications as originally conceived were to identify options for commonality and uniqueness of forms and items for different kinds of programs. These program types were defined as a first approximation to facilitate this; a first draft of the specifications included references to options. As the work progressed, it became clear that spelling out these options would require a level of detail for both the program classification and the specifications that was beyond the scope of this report. The detail could be provided only after the designation of programs to be considered as potential participants in the evaluation study.

It is evident that a large proportion of items and forms to be used for traditional programs and type 1 nontraditional programs could

be common to all of them, with branching for associate and baccalaureate degree programs. The proportion of common items would be reduced with the introduction of type 2 programs, particularly if importance is attached to association of outcomes with program features. The addition of programs of type 3 would permit the use of common items and forms only for the assessment of long-term and very general outcomes; e.g., overall satisfaction with collegiate experiences, change or advancement in career related activities, or personal growth and development.

The specifications are presented with no reference to options for common and unique information collection procedures. Discussion of them does make reference to the need for core questions and program-specific questions for certain types of information, and to the need for basing analyses on subgroups chosen on the basis of particular kinds of program experiences. As programs are identified for inclusion in the evaluation study and instruments developed, the problems inherent in the need for common instrumentation if program comparisons are to be made will loom increasingly large.

The final result of all these considerations was a division of the total evaluation study into six substudies, defined primarily on the basis of groups from whom information should be obtained, and secondarily on the basis of student degree objectives and stage of progress at the start of the proposed study. Stage of progress toward the degree objectives was used because of the two-year limitation of the length of the proposed study.

The six substudies and related questions for each are:

Substudy A: Applicants, Dropouts, and Transfers. Who are the interested learners? Why did they enroll in the program or institution? What kinds of students do not complete programs, and why? What happens to these students?

Substudies B and C: Students Working Toward or Receiving Associate Degrees (B), and *Upper Division Students Working Toward or Receiving Baccalaureate Degrees (C)*. Does the program satisfy learner needs or professional and personal interests? Does the content of the curriculum address the need of these students? Are the procedures for assessment of learner competencies and achievement creditable and valid? What are the private costs to the students?

Substudy D: Recent Degree Recipients. What happens to students after they complete their programs? How has their education contributed to educational and occupational success and satisfaction? How well satisfied are the students with what they have learned in college? What do graduates think of the program and experiences they had in college?

Substudy E: Faculty and Administrators' Perceptions of Programs. Will other institutions accept at face value the degrees and credits awarded by the programs? How credible is the program to professional personnel within and outside the program? Are there special faculty requirements or incentives? What impact have these programs had on higher education practices?

Substudy F: Program Costs and Sources of Funds. How much do the programs cost? Will the programs be able to survive without special federal, state, or foundation support? Can new programs compete with existing programs once the initial development and tryout period is completed?

It seems evident that the total study should be broken down into some set of substudies. It is not at all evident that this arrangement is the best. It was chosen after trying out several others. Given the lack of specificity concerning the purposes and context in which a multiprogram evaluation study would actually be conducted, the number of possible study designs to which information specifications could be addressed is very large. New reasonably realistic possibilities were generated almost daily as important considerations emerged. It became evident that there is no "best" arrangement. The choice of this one was made because it: a) matches the degree structure of undergraduate programs; b) recognizes the two kinds of degree objectives dealt with in the initial proposal; c) permits some comparisons among traditional and nontraditional programs, where these are appropriate; and d) allows for the presentation of information specifications for essentially intact substudies, if a comprehensive evaluation study could not be undertaken.

INFORMATION ANALYSIS PRINCIPLES

The information specifications for each of these substudies, presented in the next section, are grouped to suggest analyses that

would be appropriate to the kinds of student-program interactions described in Figure 1. It is assumed that when programs are selected for the evaluation study, some groupings of programs will be appropriate for selected analyses. While no analysis plans are suggested in this report, the following principles for data analysis and choice of respondent groups evolved as the information specifications were refined.

1. All analyses would have to be done separately for each program, and many analyses could be carried out only at the program level.

2. Some analyses might be made by combining information for different programs in the same program type, but only if the responses were not related to features of a particular program. For example, data on overall satisfaction with the college or program experience might be combined across programs; data on satisfaction with particular arrangements of faculty, materials, or resource features should not be.

3. When data are combined across programs, adjustments will have to be made for differences in the characteristics of students entering the several programs. The forms in which data on the pre-program characteristics will be available will be determined to a large extent by each program or institution. However, some scheme for deriving approximations to common forms for these indicators will have to be developed. This is discussed further in Section III.

4. If data are to be pooled across programs, or comparisons are to be made between programs or program types, crucial decisions involving technical, statistical, and measurement issues will be required. The specifications have not been refined to the level necessary to provide options related to these technical considerations. Final interpretations and conclusions will have to be heavily weighted by informal judgment.

5. The specifications generally imply data collection instruments and procedures appropriate for large-scale, complex studies. Such instruments do not presently exist in finished form, although many of them can be assembled by adaptation of existing materials. Assembly or initial development, if necessary, should be based on an initial round of rather intensive, small sample interviewing and observation. Nothing is included in the specifications on the information that should be obtained in this preliminary instrument development effort.

6. Assuming that satisfactory instruments for a large-scale study are developed as suggested above, it would still be desirable to collect additional interview or other kinds of intensive information for small samples of students during the study itself. There are too many important characteristics of programs and student groups that should affect data interpretation, but will be unknown to an external agency during the planning and instrument development stage. However well developed the large-scale instruments, it cannot be assumed that they are equally appropriate and meaningful for all programs and groups of people.

7. The measurement of pretest to posttest growth, while theoretically desirable, is generally not possible. Virtually all knowledge and skills measures are designed to provide good indicators of status at a particular time, and are quite insensitive to changes over time. In most instances, they are not even very useful as indicators of group changes unless the time interval is very long or the change is quite dramatic. If the change is, in fact, dramatic, then such a low floor and high ceiling would be required for the measure that the pre- and posttests probably would measure quite different aspects of the knowledge and skills.

The alternative of interpreting posttest information in terms of its content and in the light of different but relevant initial status information seems more appropriate. In some situations, such interpretations might be amplified by comparisons of groups using various statistical techniques to allow for differences in initial status. The use of statistical tests in such interpretations would not be appropriate, however, since random assignment of students to the different groups is not possible.

8. The specifications do not include the kinds of useful information that might be obtained using ethnographic, anthropological, or sociological methodology. This is so only because the use of such methods appears to be quite impractical within the limits of time and funds that are likely to be available for the evaluation study to which the specifications are addressed.

III. INFORMATION SPECIFICATIONS

The translation of decisionmakers' information needs and priorities to specifications for the instruments and procedures to obtain the information requires a shift in focus. The needs are generally not differentiated and the statements of them employ quite global and undefined terms. This is to be expected. Decisionmakers may use information, but they are less likely to define the needs in a way that will permit collection of information from the primary sources.

Information is obtained from people who function in particular roles in a given context. These people are students in Program A, or alumni of Program B, or faculty in Program C, or registrars at Institution D, and so on. A given information need of a decisionmaker may cut across all these roles. Where it does, it could have markedly different implications for approaches to soliciting information from the potential respondents. In effect, the information specifications resulted from a detailed consideration of these implications.

ORGANIZATION OF SPECIFICATIONS

One important purpose in the preparation and presentation of information specifications is to approach a more operational definition of the global terms used in decisionmakers' needs statements. The ultimate operational definition of these terms will be in the instruments to be used to obtain the information from the various respondents. It would be premature to attempt to provide the voluminous and detailed specifications needed to prepare specific instruments for use in given

settings, since they depend on too many planning decisions about the actual evaluation study that cannot be anticipated.

The intermediate position between specifications using only conceptual or very loosely defined terms and overly detailed instrument-specific terms has been sought. The six exhibits identify information types (content), respondent groups, possible instruments, and forms of response. These exhibits represent a first level of refinement of the global terms employed in information needs statements. The information types in each exhibit are grouped according to the elements of the model described in Figure 1. Following each exhibit is a discussion of some major issues and options for instrument development and data analysis that should be considered as a study is planned.

A second important purpose of the preparation of the specifications was the conversion of information needs priorities to priorities for potential instruments and respondents. As mentioned above, a given high priority need can cut across a number of respondent groups, and might be met using a number of different instruments singly or in combinations.

For example, the high priority information on end of program knowledge and skills could be obtained using tests and/or subsequent progress and/or employer assessment and/or faculty assessment and/or student self-assessment. Knowledgeable people who all agreed on the priority of the information need would undoubtedly differ in the priorities they would assign to each of these indicators. Yet in planning the study, some one person or group will have to determine that some instruments and procedures would be developed and used and others would not.

Suggested priorities for the content areas are included in the exhibits. They reflect considerations of information needs priorities, feasibility of instrument development, potential tradeoffs between costs and data quality and value, and adequate data interpretation. This last consideration was given considerable weight. A high priority is suggested for some content areas, not because decisionmakers asked

for the information, but because it seemed essential to the best use and interpretation of information requested. It is important that the plan for collection of information minimize as much as possible potential misinterpretations that can result from the necessary groupings and oversimplifications discussed in the first part of Section II.

Finally, the six exhibits are directed to the substudy options listed in Section II.

DEFINITION OF TERMS IN THE EXHIBITS

Abbreviated terms are used to limit the length of the exhibits. Following are definitions of the terms that are not obvious, with more extended discussion in a few instances.

Time of Collection

Start of study, periodically during study, and end of study. These terms are used relative to the schedule for the information collection phase of the evaluation study. The actual schedule for data collection will vary from program to program to be consistent with the schedules of each of the programs. Latitude in the meaning of start and end of study will be required, but it will be necessary to set some limits on them if the data collection task is to be manageable.

The collection of preprogram student data will have to continue through much of the term of the evaluation study. Some nontraditional programs permit a student to initiate studies at any time, with new students enrolling weekly or monthly. Data for these new students will have to be accumulated as students apply. In order to maintain some comparability between traditional and nontraditional programs with respect to the length of residence, probably it will be necessary to collect data at the beginning of each term for the other programs. In planning a specific study, some arbitrary limit for the collection of information for new students will have to be set.

Postprogram information for a large majority of the students in the study would be collected at the end of the information collection phase. However, such data for students who complete their degree requirements during this phase might best be collected periodically to minimize the costs of student follow-up.

Application or entrance. Some of the preprogram information can be obtained from all applicants using application forms, but it is likely that much of it can be obtained only when or after the students enroll. This too will vary from program to program. Ideally, program staff will be willing to obtain as much of the data needed for the evaluation study as possible at the time of application.

At termination. Information on the reasons for early termination (dropouts and transfers) should be obtained as soon as the student's decision to leave is known. Collection of this information should probably continue until almost the end of the data collection phase of the evaluation study.

Information Sources

Applicants. Those potential students who have completed the first form required for admission. Programs will vary in the precision with which they define and apply admissions requirements and identify applicants. Some arbitrary definitions of applicants may have to be agreed upon with the staffs of each program.

Enrolled students. Those students who have completed one or more of the steps required after the program has determined eligibility and placement within the program. Programs will vary in the ways these steps are defined, and separate definitions may be required for each program. Simple registration for courses or first meeting with a mentor or adviser, or attendance at a first orientation session, may not be sufficient to identify a student as enrolling.

Nonenrolling applicants. Any person who has completed at least an application form for a program but who does not subsequently satisfy whatever requirements are used to identify enrolled students in each of the programs.

College records. This generic term covers all institutional, research, and faculty files containing nonconfidential data about student characteristics, qualifications, achievement, and status. Programs will vary in the amount of information that can be obtained or made available from existing files. Negotiations with cooperating institutions may be required to encourage them to collect and provide as much student information as possible in order to reduce the amount of student time required for this special evaluation study.

Program staff. Personnel who serve in administrative roles or offices, as opposed to instructional roles.

Faculty/adjunct faculty. People who work with students primarily in an instructional capacity, such as teachers, mentors, counselors, orientation or workshop leaders, and resource specialists. Adjunct faculty includes those persons whose primary employment is outside of the program being evaluated but who have some paid or voluntary part-time role within the general program structure.

Possible Instruments

Coding forms. Forms used for transcription of data in college records, e.g., transcripts, admissions files, faculty ratings.

Questionnaires. Forms used to obtain responses to a set of discrete questions, most often posed so that answers can be given in defined response options.

Interviews. Structured or semistructured person-to-person discussion directed to particular topics.

Opinion scales. Forms that request verbal expressions of attitudes which by definition are self-reports. May be a section of a questionnaire or interview schedule, or may be used alone.

Rating scales. Forms for recording assessments of observable performance. May be either self-ratings or ratings of others, and may be a section of a questionnaire or interview schedule or may be used alone.

Standardized tests. Tests of cognitive functions with precise directions for administration and scoring, and with exact or suggested time limits. These tests will most often be multiple choice, but some standardized tests include items that require short or free response answers. Normative data for defined groups may or may not be available for use in score interpretation. Most standardized tests are intended principally for individual assessment; when they are, they are sometimes referred to as narrow-band instruments.

Survey tests. Tests of cognitive functions constructed for use in group assessment, using either multiple choice or free response formats. They would usually be considered to be broad-band instruments. These tests would most often be prepared specifically for this project, and normative data would rarely be available.

Reports, minutes, and other documents. These include all documents provided by program staff as sources of information in lieu of detailed interview responses.

Forms of Response for Analysis

Frequencies in defined categories. Numbers of respondents marking particular response categories. The categories may be commonly used ones (e.g., male-female, age ranges, job classifications) or ones defined specifically for this study (e.g., reasons for early termination, educational objectives); they will most often be defined on coding forms or questionnaires.

Numbers, counts, or percentages. These are essentially the same as the frequencies defined above, except that the data would be provided directly by program staff or taken from reports in which the categories were defined by the program staff and procedures.

Interview response classifications. Interview responses assigned to categories defined either before interviewing was initiated or after all interviews were completed and reviewed for popularity of various kinds of responses.

Interview summaries or excerpts. Summaries of, or excerpts from, interview information that cannot be systematically classified. These will most often be used when there is a limited number of a particular type of interviewee, or to provide illustrative support for conclusions based on more restricted instruments. Excerpts might also be used if information from an interview could not be assigned to an interview response category, but was considered to be quite important.

Single opinion scores. Numerical value of points in a continuum of reactions to a statement. The continuum may be defined by: (a) subjective value labels for the end points (e.g., useful/not useful); (b) subjective value labels for end and intermediate points (e.g., very useful, useful, no opinion, etc.); or (c) actual or implied scale values (e.g., useful for both personal development and job preparation, useful for personal development, useful for job preparation, etc.).

Summed opinion scores. Sums of scores for responses to a set of opinion items judged to be projections of a common attitude domain. The scores that are summed may be: (a) arbitrarily determined (five-point continuum assigned values 1-5), (b) empirically determined by judges (Thurstone scales), or (c) derived from internally consistent responses (Likert scale).

Single rating scores. Numerical values on a continuum defined by: (a) behavioral descriptions of the end points (e.g., can present complicated ideas in a very understandable way/cannot communicate simple ideas effectively); (b) behavioral descriptions of end and intermediate points (e.g., can write a publishable research report, can write an acceptable term paper, can write an acceptable letter of application, etc.); or (c) normative statements (e.g., top 10%, between 60th and 90th percentiles, between 40th and 60th percentiles, etc.).

Summed rating scores. Sums of scores for ratings of individual characteristics which in aggregate define a single more general variable.

Test scores. Numbers on a scale of cognitive functioning which has at least ordinal meaning, with demonstrable reliability adequate for at least the differentiation of definably different groups.

EXHIBIT A

Information Specifications for Substudy A
Applicants, Nonenrolling Applicants, Dropouts, and Transfers

Purposes of information collection

1. Description of interested learners. (Who are the interested learners? Why did they enroll or not enroll in the program or institution?)
2. Study of dropouts, transfers, etc. (What kinds of students do not complete programs, and why? What happens to these students?)
3. Student grouping or baseline data for subsequent analyses in other substudies

Primary users of information

1. Program staff, institutional and state planners
2. Program staff, institutional and state planners
3. Evaluation study staff

Information type	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of responses for analysis	Suggested priority
A. Students					
1. Sex, age, and ethnic background	Application or entrance	College records and/or Applicants	Coding form and/or question-naire	Frequencies in defined categories	High
2. Preprogram educational background a. Secondary school b. College degree directed c. College credit courses--no degree objectives d. Noncredit, college sponsored education e. Other					
3. Educational objectives a. Highest degree b. Current degree c. General educational & career goals					
4. Current occupation					
5. Preprogram occupations					
6. Occupational objectives a. Specific to immediate degree program b. When current degree program is completed c. Completion of ultimate degree		Applicants	Question-naire		
7. Academic ability a. Verbal b. Quantitative c. Other		College records or Faculty and program staff or Applicants	Coding forms or Rating scales or Self-rating scales	Standardized test scores from admissions/guidance/orienta-tion files or Single rating scores based on normative statements (A)	
d. Overall				Summed rating scores (A)	

EXHIBIT A (continued)

Information type	Time of collection	Information sources	Possible instruments	Preferred (P) and Acceptable (A) forms of responses for analysis	Suggested priority
A. Students (continued)					
8. Knowledge a. Humanities and literature b. Mathematics and natural science c. Social sciences	Application or entrance	Applicants	Standardized tests or Survey tests or Self-rating scales	Test scores in each area on comparable scales (P) or Broad band test scores on comparable scales (A) or Single rating scores based on normatively defined points (A)	High
d. Overall				Summed rating scores (A)	
9. Skills a. Written communication (reading, writing) b. Oral communication (speaking, listening) c. Vocational/occupational d. Other (artistic, musical, mechanical)					
B. Program procedures					
1. Needs & expectations for program arrangements a. Schedule for classes, seminars, studies b. Location for group meetings c. Methods of learning, contact with faculty d. Content of studies e. Financial assistance f. Information, guidance g. Orientation for reentry h. Certification for learning			Opinion scales	Single opinion scores based on implied scale values (P) or Subjective labels for end and intermediate points (A) and Summed opinion scores for subsets of related needs (P)	
2. Reasons for decision not to enroll	Periodically during study	Nonenrolling applicants	Interview or Questionnaire	Interview classification (P) or Frequencies in defined categories (A)	
3. Reasons for early termination	At termination	Students terminating without a degree			
4. Assessment of programs and procedures a. Schedule for classes, seminars, studies b. Location for group meetings c. Methods of learning, contact with faculty d. Content of studies e. Financial assistance f. Information, guidance g. Orientation for reentry h. Certification for learning i. Overall			Opinion scales	Single opinion scores based on implied scale values (P) or Subjective labels for end and intermediate points (A) Summed opinion scores for subsets of related needs (P)	

EXHIBIT A (continued)

Information type	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of responses for analysis	Suggested priority
C. Outcomes					
1. Numbers of applicants who do not enroll	Periodically during study	College records	Coding form	Frequencies in defined categories	High
2. Numbers of students terminating enrollment prior to receipt of degree a. No plans for continued education b. Transferring to other institutions/programs c. Reenrolled in same program d. Units completed, area of study e. Achievement (GPA if applicable)	End of study				
3. Faculty ratings of students leaving without degrees a. Ability b. Achievement (including GPA) c. Interest d. Objectives e. Financial and other problems f. Overall	At time of termination	Faculty	Rating scales	Summed ratings or Single ratings with defined points	Medium to low Medium
4. Educational activities a. Degree work, same field as in program b. Degree work, different major field c. Credit courses, no degree plans d. Occasional noncredit courses or activities e. No educational activity	End of study	Nonenrolling applicants and Students terminating without degree	Questionnair and/or Telephone interview schedule	Frequency in defined categories	High
5. Educational plans a. Short-term b. Ultimate					
6. Occupational activities a. Current employment b. Anticipated change in employment, if any or c. Reasons for unemployment					
7. Occupational plans a. Short-term b. Ultimate or career directed					

EXHIBIT A (continued)

Information type	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of re-sponses for analysis	Suggested priority
D. Costs					
1. Financial obligations	At termination	Students terminating without degree	Rating scales	Single rating scales based on relevant dollar ranges (P) or Normative verbal descriptions(A) and Summed rating scores	High to medium
a. For educational costs					
b. Personal/family/dependent support					
c. Other					
d. Total					
2. Financial resources					
a. Savings					
b. Support from parents and/or spouse					
c. Employer contribution					
d. Concurrent employment					
e. Scholarships, loans, federal support					
f. Total					

ISSUES AND OPTIONS FOR SUBSTUDY A INSTRUMENT DEVELOPMENT

Two major questions will require answers in planning for the collection of information outlined in Exhibit A: Should the reasons that some interested applicants do not enroll be determined? Should the subsequent educational and occupational experience of these nonenrolling applicants be documented?

Nonenrolling Applicants Information and Cost Tradeoffs

Nonenrolling applicants are presumably one segment of the program market--a very important segment if programs are seeking potential modifications to meet previously unmet needs. Additionally, knowing why some of the interested applicants did not enroll would provide data that could be very useful in understanding the reasons other interested applicants do enroll. Thus, the first question is closely related to the first purpose for Substudy A.

The second of these two questions in effect extends the definition of dropouts to include anyone who applies for admission but does not ultimately receive a degree from the program. A two-year follow-up of nonenrolling applicants could add appreciably to the information useful in contrasting the educational and occupational progress of those who did and did not complete the degree program.

The cost of data collection and analysis for nonenrolling applicants will be high, however, because they will present many "missing data" problems. Some programs prefer to keep application forms and procedures as simple as possible, using an initial preenrollment orientation period as a means to a mutual (program and student) decision about enrollment. If an applicant decides not to complete the orientation, very limited data--perhaps no more than name, address, telephone number, and years of high school completed--would be available. If an applicant decides after completion of the orientation session that a different program would be more useful, only information collected during the orientation might be available--such things as orientation worksheets or evaluation forms--but little more. These orientation

byproducts may or may not include such things as demographic data, objectives, or needs and interests. In other instances, students may be asked to indicate their decisions in writing, with reasons for the decision.

Other programs may have a sequence of application steps--completion of an application form, submission of admission test scores, personal interviews, and so on. Applicant information is obtained at each of these steps, but applicants may discontinue the process at any step, so that the information is in varying stages of completeness.

Still other programs may admit students fully on the basis of some amount of applicant information, but require specified first steps--preparation of a portfolio, completion of noncredit skills tutorials--prior to enrollment for instruction. Students are "enrolled" in these first steps, but those who drop out prior to enrollment in instruction in a sense were never enrolled. Again, varying amounts of information will be available.

Each of these examples suggests different kinds of missing data problems that would arise even if Substudy A were limited to the reasons for nonenrollment. They also suggest a number of definitional problems that will have to be resolved. These problems will be compounded if the substudy is to include the two-year follow-up. Most of the nonenrolling applicants will have had little contact with the program during this time, so that locating them and making initial contact with them will be time consuming and expensive. Once contacted, they will probably have very little interest in responding to requests for information for an external evaluation study. In fact, those who do respond to such requests may represent a sample that is biased in important and unknown ways, but the paucity of information from the time of application will make it very difficult to detect or allow for such biases.

In summary, inclusion of nonenrolling applicants in a multiprogram evaluation study would provide valuable information but would greatly increase the complexity and costs of the study. The potential value

and quality of the information will have to be weighed against the added cost and time for this aspect of the substudy. Omission of one or both of the questions could reduce the cost of the total evaluation study, or could allow for Substudies B and C to be done more thoroughly.

Definition of Dropouts and Transfers

As increasing numbers of students attend college on a parttime basis, often combined with employment, the difficulty of identifying dropouts and transfers increases. This may be particularly true for nontraditional programs, many of which are designed to give students great latitude in determining the rate of learning progress. Students may "stopout" for a time with the full intention of returning; even though they may not be enrolled for long periods they cannot be classed as dropouts. In other cases, programs may permit or encourage students to take courses at another institution in order to obtain a more satisfactory sequence of studies; such students cannot be considered transfers.

Definitions will have to be negotiated on a program-by-program basis. In some cases it will almost certainly be necessary to make an initial contact with students not currently enrolled for the purpose of determining whether they have dropped out or have transferred to another program. Some rather arbitrary definitions may ultimately be required.

A two-year study may provide only partial answers to the important questions about dropouts and transfers for parttime adult students. A useful supplementary study to find out what happens to these students could be done using all students who had been enrolled in the year prior to the start of the evaluation study but did not enroll during the term of the evaluation study. These students could be contacted at the end of the evaluation study to learn what had happened to them. Although some of the applicant information collected at the start of the study might not be available for these students, this lack would be more than offset by the information that had been accumulated while

they were enrolled in the program (e.g., length of time in the program, number and length of prior nonenrollment periods, grades, etc.)

Information Type and Response Forms

Some of the items listed under "students" may be available from presently used application or personnel forms. It would be most desirable, however, if data could be collected from all programs in a generally uniform way. This will require a survey of the application forms for all participating programs, followed by negotiation with each program on forms and procedures that will minimize required variations across all the programs.

Educational and occupational experiences and objectives. Items A-2 through 6 are particularly important information for some nontraditional programs. While they should also be viewed as relevant to student success and satisfaction in traditional programs, such programs may not attach as much importance to them. Certainly, they may not see a need for obtaining the amount of detail thought to be important by some of the participating nontraditional programs. Some compromises will be necessary if reasonably common information is to obtained for all programs. The most likely solution may be one of having a common core of questions or information items that could be asked of students in all programs, with additional detailed questions that will vary for different programs. It will be important to accommodate the wishes of each participating program staff in order to gain their cooperation, and to increase the value of the study results to the program; this could result in separate data collection instruments and formats for many programs. If this is the case, complex administrative and data management problems will be created at a later stage of the study. The tradeoffs between collecting detailed information appropriate to each program or a common core of more general information will have to be weighed frequently and as objectively as possible.

Information about prior educational experiences should be substantially different for those students seeking baccalaureate degrees from that for students seeking only associate degrees. Secondary school experiences will be less important for the former, but may be the only formal educational experiences for the latter. On the other hand, the nature and variety of collegiate-level educational experiences would be of great importance for upper-division baccalaureate students (Substudy C) but may be difficult to obtain and evaluate in a common frame of reference. In any case, it would be of great value to obtain for all students some quantitative assessment of the quality of performance in the prior educational settings, if some procedures can be developed for placing these estimates on an approximately common scale. At the very least, where transcripts of prior work are available, best estimates of grade-point averages should be obtained or calculated using the same arithmetic procedures for all students and programs.

Written and oral communications skills. Information types A-9, a and b are high in priority for two reasons. First, many decision-makers indicated level of achievement in these skills as a high priority measure of program effectiveness; it is important, therefore, that some assessment of start-of-program status be obtained. Second, these entry skills--particularly the written communication skills--are undoubtedly closely related to subsequent student success and satisfaction in the program. Some nontraditional programs have directed particular attention to the development of these skills in order to increase access to higher education for some nontraditional students. Many traditional programs have also initiated skill development efforts for various student subgroups, but the nature of and context for the two kinds of program efforts (nontraditional and traditional) may be quite different.

The range of skill levels that will likely be encountered in a multiprogram study, combined with a desired focus on college-level skills, will make it difficult to choose a single test. Several

publishers now have tests intended for college-level use that may be appropriate, particularly for the assessment of reading skills. These tests may, however, require more time than necessary if the only purpose of the information is group assessment and statistical adjustment. In addition, the content may be judged to be inappropriate for older adults who have not been recently enrolled in secondary or college classes. Increasing numbers of colleges and universities are preparing their own writing skills tests. The primary purpose of these tests is individual student placement, however; again the tests may require an inordinate amount of time if group assessment is the sole purpose for collecting the information. It may be that relatively short survey tests of reading and writing should be prepared specifically for this study. The range may have to extend from the level of adult "survival" skills tests to the level of scholastic aptitude college admissions tests.

There apparently are no readily available standardized oral communications skills tests; scales developed specifically for this study will be required if these skills are to be included. In fact, the most feasible approach to assessing all four skills may be that of using specially developed rating scales with functionally defined points. Use of such scales would be limited to enrolled students rated by faculty or other staff at the end of, say, a six-week period of enrollment.

Needs and expectations for program arrangements. Information type B-1 is particularly important for an evaluation of nontraditional programs. Major features of many such programs stress responses to special needs of parttime adult students, including flexibility of schedule and location for learning, credit award for prior experiential learning, and various kinds of student support services. Some traditional programs have also made various kinds of responses to these needs as increasing numbers of students of all ages are undertaking parttime study combined with work. Regardless of program design, however, students may enter both kinds of programs with expectations that could affect their ultimate assessment of the program features.

It will probably not be sufficient simply to ask applicants whether they have some of these needs, or whether the needs are "Very great, Great, Of little importance, or Nonexistent," for example. Such responses would be entirely defined by each student's personal perception of need. Without some reference to program-related options, they may not be useful for grouping students for subsequent analyses. Ideally, the items or questions might determine the strength of the needs by having applicants indicate identified alternatives, which could be ranked by judges in terms of presumed strength of needs. A possible example item, which has not been tried out, is:

> *If a program did not permit me to have personal contact with faculty members on a regular basis, I would*
>
> *1. not continue beyond the first semester*
>
> *2. look for activities outside of class that would allow for such faculty contacts*
>
> *3. seek such contacts with other program personnel in counseling services and so on*
>
> *4. become involved in extracurricular activities with other students*
>
> *5. spend more time on course-related reading and study.*

The options for this example are intended to indicate varying strengths of need as reflected by the actions people say they would take if the need were not met by the program. The first two options both suggest a very strong felt need for personal contact with faculty, but the first one indicates a stronger need because the action is more drastic. The third option suggests a need for contacts with professional staff, but the possible action indicates that it is not necessarily a need for intellectual stimulation. The last two options indicate very little need for faculty contact; they differentiate strength of need for some kind of personal contact through the program, since option 5 is intended to indicate virtually no such need.

This item is clearly not in a final form for pretesting. The example is presented only to illustrate a kind of effort that is needed to move from reliance on sets of independent, very general opinion

and rating items toward the use of scales that justify more comprehensive analyses of the interactions between students and program features. If a number of similar items relating to student needs could be written and scaled through pretesting, total and subscores could be derived. Such scores might have measurement properties that would indicate more accurately the strengths of various needs of different kinds of students. The development and pretesting of such measures will require extensive time and effort if all the needs listed for information type B-1 are to be assessed.

Assessment of program and procedures. Information type B-4 is virtually identical to B-1. The items and instruments used for assessment should probably be different, however, since the students have had additional relevant experiences between the two times for information collection. It seems likely that at least some of the items for B-4 should be specific to the features of the particular program in which the students were enrolled. Whether items could also be written in terms that would be general to all programs is not at all clear.

An example of an item that might be tried out for information type B-4 is given below. The intention in writing the example was to construct an item related to the intent of the previous example so that analyses of program assessment data might be related in meaningful ways to preprogram indicators of student needs.

> *The opportunities for personal contact with faculty were:*
>
> 1. *so infrequent that this was an important factor in my decision to leave the program*
> 2. *infrequent in formal instruction, but I was able to find other settings where faculty contacts could be made*
> 3. *of little importance because I was able to make personal contact with staff in other settings*
> 4. *of little importance because of the value of my contacts with other students*

5. *of little importance because I learned what I wanted to on my own*
6. *quite frequent and very helpful to me.*

The development of program assessment items will require even more thought and creativity both before and after pretesting than is the case for student needs items. A large number of items similar to the example will be required for tryout to obtain internally consistent total or subscores. There are so many important dimensions to the interaction of students with programs that writing unambiguous items will present a real challenge. If reliable assessment of the responsiveness of educational programs to student needs is to be achieved, this extended effort would appear to be necessary.

In discussing above the desirability for collecting information for a number of different programs using reasonably common forms, mention was made of the possible need for having some core questions common to all programs, and other questions specific to each program. It seems certain that this procedure will have to be followed in collecting data for information types B-1 and -4. Data would misrepresent a program seriously if they were obtained using only items that are the same for all programs. The data will limit severely the generality of any conclusions if the items and forms are unique for each program.

Faculty ratings of students leaving without degrees. Information item C-3 is included in the specifications because it would be useful to have the faculty perspective on the characteristics of dropout and transfer students as a supplement to the information provided by the students. The data would be useful in corroborating or enlarging on the reasons for early terminations. They may also be useful in analyzing the information obtained from these students at the end of the study regarding their experiences and progress after leaving the program.

Identification of faculty and staff who can provide such information based on sufficient acquaintance with the students may pose problems,

particularly in larger programs. It will be essential that the information be obtained as soon as possible after the student makes the decision to leave, which will be a burden for the faculty--perhaps even a frequently repeated annoyance. In many instances, the students may not have been in the program long enough for any faculty members to know them well enough to assess the factors listed in Exhibit A. In other cases, the detailed information listed may require that different faculty be asked to provide different bits of the data. It is possible that useful information will simply not be obtainable for appreciable numbers of students regardless of the effort expended, and that those for whom it is available will be very biased samples of the dropouts and transfers. The suggested priority for this information type is "medium," not because it is less important than the data provided by the students, but because of the assumed difficulty of obtaining it. There will be an important tradeoff between data quality and usefulness in relation to cost and effort to get the information.

End of study information for nonenrolling applicants, dropouts, and transfer students. The general terms for the types of outcomes information to be collected at the end of Substudy A are similar to the terms used at the end of Substudies B, C, and D. Less detail in the response forms is suggested as one of the ways to increase the rate of cooperation. The end-of-study information for these special subgroups should be useful in part to compare their outcomes with those for students who receive degrees. The ease of interpretation of such comparisons will be very closely related to the degree to which the end-of-program data for the several groups can be collected using similar forms.

EXHIBIT B

Information Specifications for Substudy B
Students Entering Program with Associate Degree Objectives

Purposes of information collection	Primary users of information
1. Assess outcomes, accomplishments, and satisfaction (Does the program satisfy learner needs or professional and personal interests of students interested in associate degrees?)	1. Institutional presidents and deans, state planners and legislators, program administrators
2. Student satisfaction with program procedures (Does the content of the curriculum address the needs of the learners? Are the procedures for assessment of learner competencies & achievement creditable & valid?)	2. Institutional deans, program administrators, faculty, and staff
3. Indirect costs to students (What are the private costs to students?)	3. All decision groups

Information type	Time of collection	Information sources	Possible instruments	Preferred (P) and Acceptable (A) forms of responses for analysis	Suggested priority
A. Students					
1-9. Preprogram data listed in Exhibit A					
10. Credit transferred for prior formal education a. Amount b. Level of performance in prior courses	Application or entrance	College	Coding form	Amount by area of study & total, with estimates of performance level	High
11. Credit awarded for prior experiential learning a. Amount b. Judged level of performance, if applicable					
12. Degree requirements waived for prior educational or experiential learning					
B. Program procedures					
1. Preprogram data listed in Exhibit A (Needs & expectations for program arrangements)					
2. Assessment of program and procedures a. Schedule for classes, seminars, studies b. Location for group meetings c. Methods of learning, contact with faculty d. Content of studies e. Financial assistance f. Information, guidance g. Orientation for reentry h. Certification for learning i. Overall	End of study*	Students	Opinion scales	Single opinion scales based on implied scale values (P) or Subjective labels for end and intermediate points (A) Summed opinion scores (P) or single opinion score with implied scale values (A)	High
C. Outcomes					
1. Current educational objectives a. Change in current degree b. Change in ultimate degree c. Change in other			Questionnaire	Frequencies in defined categories	
2. Current occupational objectives a. When receive degree b. Ultimately					

43

EXHIBIT B (continued)

Information type	Time of collection	Information sources	Possible instruments	Preferred (P) and Acceptable (A) forms of responses for analysis	Suggested priority
C. Outcomes (continued)					
3. Student assessment of progress toward original and/or revised educational and occupational career objectives a. Academic skills b. Occupational skills c. Knowledge d. Personal development e. Clarification of goals and objectives f. Overall	End of study*	Students	Questionnaire or or Rating scales	Frequencies in defined categories (A) or Single or multiple self-rating scores in each area	Medium to high
				Summed ratings scores (P) or single rating score with defined end and intermediate points (A)	High
4. Collegiate level written communication skills a. Reading b. Writing			Standardized tests or Survey tests	Scores over a wide range of ability levels	
5. Adult level oral communication skills a. Speaking b. Listening		Faculty or other staff	Rating scales	Ratings with functionally defined points (P)	Medium
6. Assessment of occupational skills development		Faculty/ adjunct faculty	Rating form	Single rating scores with functionally defined points (P) or with subjectively defined end and intermediate points (A)	Medium to low
7. Assessment of knowledge of broad areas of study a. Humanities and literature b. Mathematics and natural science c. Social sciences		Students	Standardized tests or Survey tests	Test scores in each area on comparable scales (P) or Broad band test scores on comparable scales (A)	Medium Medium
		or Faculty in each area	or Rating scales	or Single or multiple scores in each area based on defined end and intermediate points (A)	Medium to low
8. Degrees completed		College	Coding form	Counts, numbers, & percentages	High
9. Progress toward degree objectives (for students still enrolled)					

* Students receiving degrees during the term of the evaluation study might best be queried at the time the degree requirements have been completed.

EXHIBIT B (continued)

Information type	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of responses for analysis	Suggested priority
D. Costs					
1. Private costs to students and other financial obligations a. Lost job time b. Transportation c. Babysitters, etc. d. Relocation e. Duplication of housing f. Family/dependent support g. Other	End of study*	Students	Questionnaire and/or Rating scales	Frequencies in defined categories and/or Single rating scores based on relevant dollar ranges (P) or normative verbal descriptions (A)	High
2. Financial resources a. Savings b. Support from parents and/or spouse c. Employer contribution d. Concurrent employment e. Scholarships, loans, federal support f. Total				Summed rating scores	

* Students receiving degrees during the term of the evaluation study might best be queried at the time the degree requirements have been completed.

ISSUES AND OPTIONS FOR SUBSTUDY B INSTRUMENT DEVELOPMENT

This substudy is limited to students working for associate degrees because such programs differ from baccalaureate programs in nature and expected outcomes. As the evaluation study progresses, students identified for inclusion in Substudy B at the start of the study will separate themselves into at least four subgroups. Some will drop out or transfer to other programs; when they do, they will become subgroups in Substudy A. Others who receive substantial amounts of credit at entrance for prior learning and experience will complete all the degree requirements at various times during the term of the study. Most full-time students will complete the requirements at the end of the term of the evaluation study, while parttime students will still be pursuing their degree work at this point.

The term "students" has been used in Exhibit B to include the last three of these four subgroups. The same general kinds of information should be obtained for all these students, but some of the details should vary to reflect the different stages of progress toward the degree. Information for students who complete their degree requirements appreciably before the formal end-of-study data collection period will have to be contacted when they complete their degree requirements to provide this information. (Although these students will become alumni during the evaluation study, few if any of them will have received their degrees early enough to meet the suggested criteria for inclusion in Substudy D, which is directed to alumni who received degrees one to three years prior to information collection.)

Issues for Student Group Definition

Substudy B is limited to students who initiate their degree work during the term of the evaluation study, in order to simplify the preprogram data collection task. The task of collecting the same or similar kinds of entry data for students in a number of different programs is too complex to rely on the use of historical data for students who are already enrolled. Even with this limitation,

there will be a number of student group identification issues that will require some arbitrary decisions if the data collection task is to be manageable.

One major issue to be considered is that many students, particularly older ones, may initiate a program by accumulating credits, with no specific degree objective. Only after they have explored and clarified their interests and personal objectives, and gotten a clear indication of their ability to handle college work successfully, will they choose a particular area of study and a degree objective. The limitation of this substudy to those students who have settled on the associate degree objective will exclude some students for whom particular nontraditional programs were designed. It was imposed because of the need to set definable boundaries on the respondent group. Without some boundary, the study could include a large number of students whose program participation was so limited and sporadic that no observable program effect should be expected. The restriction that the students have established the degree objective at the time of entry is less defensible. It was based on two considerations: (a) desirability of limiting start-of-study data collection to those students for whom there is an expectation of sufficient program participation to yield some observable program effect; and (b) simplification of data management and analysis problems by limiting the potential number of student subgroups for analysis. It would be possible to modify the second restriction to make the substudy design more consistent with the progression of parttime and/or adult learners toward degree objectives. For example, the limitation might be broadened to include students enrolled for sufficient credits that an associate degree could be obtained within, say, four years if the student chose to work for a degree and if the same load were carried throughout this time. Or it might be broadened still further to include any student who indicated matriculation for a degree at some time in the future as being probable. It should be emphasized, however, that any broadening of the limitation will result not only in an increase in the data collection and

management task, but will also complicate seriously the study of dropouts and transfers (Substudy A).

A second issue to be considered in the definition of the substudy stems from the fact that groups of students may be entering the programs, and other groups completing the degree requirements, continuously during the term of the evaluation study. This may not be an important problem for most traditional programs that have relatively few times when students may enter--fixed semesters or quarters--and have fixed residency requirements for degree award. It will, however, be a crucial problem for nontraditional programs. Some of them are designed to permit entry at any time appropriate to student needs and readiness. Others are designed with flexibility in residency requirements depending on amounts of prior formal and experiential learning for which credit is awarded.

It will be crucial to make provision for continuing collection of start- and end-of-program information at certain program sites, but this will increase the cost of data collection by more than the proportional increase in the numbers of additional students. To achieve a reasonable balance between costs and quantity of data, a time will have to be set after which collection of preprogram information will be discontinued. Another point before which no postprogram data will be obtained will also have to be set.

Information Types and Forms of Response

Some of the general issues to be dealt with in the collection of preprogram student information (types A, 1-9) were discussed in association with Exhibit A. The items and forms to be used for the collection of information in Substudy B (associate degree students) will have to be different in many respects from those to be used in Substudy C (baccalaureate degree students).

Most of the associate degree students will have had little formal education beyond high school; the quantitative data summarizing high school experience should be sufficient in most instances. In the

nontraditional programs, however, significant numbers of students will have accumulated college credits by taking courses without any degree objective, and others may have had various noncredit college courses. It will be necessary to provide for the collection of at least the amount of college credit accumulated by these students.

Ability, knowledge, and skills. There will undoubtedly be great variation among associate degree programs, both traditional and nontraditional, with respect to available measures of academic ability (A-7). Such data, if they are available at all, will ordinarily have been collected as part of a postadmission guidance and counseling program, rather than as the result of an admission requirement using one of the more common admissions tests. If test data are not available in college records, it will probably be difficult to persuade the programs to collect it in a special test administration for the purpose of this evaluation study.

Similarly, measures of knowledge in broad areas (A-8) may be quite inappropriate for many associate degree students, while vocational and occupational skills measures may be very appropriate, at least for students in occupationally oriented degree courses. Unless students are to be subgrouped for analysis by educational objectives at the level of major field, it will most likely be necessary to devise some scheme for converting various knowledge and skill indicators to some common derived scale.

The one skill area where a concerted effort should be made to obtain information of as high a quality as possible is that of written and oral communication skills (A-9, a and b). Such data will be of great importance to the interpretation of related outcomes measures that are a high priority kind of information for many decisionmakers. Comments about the needs and difficulty of assessment of the skills were made in association with Exhibit A.

Assessment of ability, knowledge, and skills at entrance will probably have to rely heavily on student self-ratings, ratings by program staff, and some scheme for converting scores from an array of

essentially noncomparable tests to a common scale. If staff ratings are used, they probably should not be obtained until at least a month after the students have been in the program, so that staff will have some opportunity to have become acquainted with the students.

Credits transferred or awarded, and degree requirements waived. The collection and organization of information about prior learning that is relevant to degree requirements will pose serious problems if data are to be pooled across programs. Programs vary widely in the ways in which they deal with the transfer and award of college credit for prior learning. This is particularly true of many nontraditional programs.

Clearly, the issue of credit for prior formal education (item A-10) is simpler than that of credit for prior experiential learning (A-11). Traditional programs have been dealing with the former for some time. However, the common currencies of credit hours and grades are becoming less meaningful as the variety of higher education instruction and assessment procedures increases. It may be necessary to devise procedures for this evaluation study to obtain some kind of common derived indicators of quantitiy and level of performance of prior education.

Award of credit for prior experiential learning is still an evolving practice. Most of the effort has necessarily been directed to procedures appropriate for individual assessment. The use of College Level Examination Program tests, the New York Regents Proficiency Tests, now available through the American College Testing Program, or procedures developed by the Cooperative Assessment of Experiential Learning project, would require too much student time and money to be practical for program evaluation. There is probably little uniformity among different programs in the procedures for award of such credit, and in fact the procedures within particular programs may vary markedly over time and for different students. The state of the art is such that perhaps the most that can be done for a program evaluation study is to group students having received or not received such credit, or to record the percent of total degree requirements that were satisfied through experiential learning.

Assessment of program procedures responsive to student needs. The discussion of instrument development for information types B-1 and 2 for Substudy A is generally applicable to Substudy B. Little more can be said without experience in the development of such information. It does seem likely, however, that more items and subscores would be required to cover adequately the longer-term involvement of those students who continue in the programs. It is also very possible that scales or scores of student assessment of progress toward their objectives (information type C-3) will be closely related to scores for B-2. The two sets of opinion indicators or ratings may simply be different ways of assessing essentially the same attitude.

Progress toward educational and occupational objectives. The focus of ratings for information type C-3 ideally should be on change over time of the factors related to student objectives. This ideal is not attainable for important reasons. Many students' objectives should and do change as the result of their continued education; for them, this may be the most important outcome. Thus, if progress toward student objectives is the focus, the times at which initial and final status should be assessed are defined only by the students. The alternative of using objectives defined by the program, and collecting data at times specified by the program, is contrary to the purposes of most educational programs. This is particularly true for nontraditional programs of the second and third types defined in Section II.

A potentially useful procedure may be to ask students for ratings of end-of-study status on at least each of the factors listed in Exhibit B in conjunction with ratings of the importance of each of these factors to the individaul student's objectives. Stated importance of the factors and self-perceived status on each of them may, of course, be confounded in important ways. Undoubtedly, a great deal of item development work and analysis will be required to know whether this approach will provide useful data for meaningful analysis.

However the ratings of end-of-study status are obtained, they should be analyzed using preprogram test scores and/or ratings on

relevant variables for statistical adjustments or for identifying subgroups of students to be contrasted on these end-of-study ratings. This will require careful consideration of detailed plans for analysis when the rating scales are developed; these analysis plans will depend, in turn, on the procedures and arrangements that can be made for the collection of preprogram data. There is, therefore, little more that can be suggested now for the development of the rating scales.

Student self-ratings are the most frequently used procedure for assessment of educational outcomes. The pool of potential items that might be used or adapted for the ratings is very large. If existing scales using functional or well defined normative descriptions of the points on the scales can be identified, they will be more useful than items that make use of subjectively defined points.

As noted above, ratings of progress may be closely related to assessment of program procedures. It will still be important to get both kinds of ratings. Many state-level decisionmakers will attach more importance to student assessment of progress resulting from collegiate enrollment than to assessment of the program procedures. On the other hand, program and institutional decisionmakers will probably be particularly interested in student assessment of program procedures, in order to see how they might improve the program.

Communication and occupational skills. The collection of data for information types C-4, 5, and 6 will be particularly important for associate degree students. The procedures employed for assessment of communications skills may have to vary for the several programs, however. For example, one program may have the development of communication skills as an objective only for those students who cannot demonstrate a predetermined level of competence at entry. A second program may adopt this as a general goal but recognize that students enter with different levels of skills, and should be expected to exit with increased but still different levels of skills. Another program or institution may accept only students who have satisfied program requirements for communication skills, and devote no explicit attention to

skill development. If a single set of instruments were to be used for all of these programs, they would have to be quite broad in range; in fact, different instruments for different levels of skill difficulty may be appropriate. Certainly the analyses and interpretation of the data should be quite different for each of the programs.

The problems in locating or developing appropriate communications skills instruments were discussed in association with Exhibit A. In a sense, these problems assume greater importance at the end of the study, however, because they are associated with outcomes indicators. Some ongoing development work by an external agency (ACT) may have progressed to a point where appropriate instruments will be available at the time end-of-study data are to be collected. If, however, only faculty ratings can be used, it will be essential that the rating scales be developed with a clearly stated normative frame of reference and/or functional definitions on the rating scales.

The assessment of occupational skills (C-6) will be important for many of the associate degree programs and students. The choice or development of instruments will undoubtedly have to be done in the context of program and course objectives. It may even be necessary to develop different instruments for several groups of students identified on the basis of student objectives (e.g., business, computer technology, graphic arts).

Survey tests for the several occupational skills might be assembled from various certification tests, or from assessment devices used by professional and occupational associations. Possible opportunities for cooperative development efforts directed to program assessment would have to be explored. If pools of skill items could not be located through such sources, a possible alternative would be to assemble course or classroom tests used in the cooperating program to provide a pool of items from which the survey tests could be assembled.

Knowledge in broad areas of study. The inclusion of information type C-7 is responsive to the expressed interest of an appreciable number of decisionmakers. Such information should probably be obtained

for all students for whom occupational skills tests are *not* appropriate. Alternatively, the group for whom this information should be obtained might be all those students whose ultimate educational objectives include the baccalaureate degree.

Even though such data could be important in assessing the outcomes of degree programs for many students, the suggested priority for the information type is "medium." This was based on considerations of the high cost in dollars and student time required to obtain reliable information, the fact that on the average decisionmakers ranked this information lower than other outcomes indicators, and the fact that it would be relevant for only some programs or student groups included in Substudy B.

Assessment of general knowledge for students enrolled in degree programs has, of course, been a problem to which extensive effort has been directed over many years. For the most part, the instruments are intended for individual student evaluation and are organized by broad or restricted discipline areas. The choice of broad areas of knowledge as the focus for Substudy B is based on the assumption that associate degree programs generally will not permit sufficient specialization to justify the use of specific discipline or subject tests. Survey tests for use in program evaluation are needed. The tests should cover a broad range of knowledge, and item sampling techniques would be most efficient.

Private costs and financial resources. Private costs of higher education to students (information type D-1) are important for all students, and may be particularly crucial for older parttime students in nontraditional programs. This information should be used in close conjunction with the program cost information collected in Substudy F, but it will probably not be possible to obtain the same level of precision. Distinctions will have to be made between lost job time and income and foregone earnings, perhaps rather broad ranges of cost categories used, and somewhat equivocal definitions of time periods for estimating costs. While private costs are a high priority item as

part of the overall cost information, it may be so difficult to obtain that the effort should be made only for programs that can also provide information for Substudy F. In other words, it seems likely that private student costs should be obtained only if good data on direct program costs can be obtained.

With the increase in enrollments of parttime students and older adults who have many kinds of obligations, it is important to understand better the many sources on which they draw to finance their education (information type D-2). Although it might be useful to know about student plans at entrance, estimates at that time might be quite unreliable; collection of data at end of study will be more useful since the estimates will be based on actual experience.

EXHIBIT C

Information Specifications for Substudy C
Students in Upper Division (or equivalent progress toward degree) with Baccalaureate Degree Objectives

Purposes of information collection	Primary users of information
1. Assess outcomes, accomplishments, and satisfaction (Does the program satisfy learner needs or professional and personal interests of students seeking baccalaureate degrees?)	1. Institutional presidents and deans, state planners and legislators, program administrators
2. Student satisfaction with program procedures (Does the content of the curriculum address the needs of the learners? Are the procedures for assessment of learner competencies and achievement creditable and valid?)	2. Institutional deans, program administrators, faculty, and staff
3. Private costs to students (What are the private costs to students?)	3. All

Note: *The types of information to be collected for students doing upper-division work will be similar in most respects to those for students starting work on associate degrees. It would be expected, however, that the wording of questions or scales to solicit this information would be markedly different in many instances, because of the differences in educational objectives and prior educational experiences. These differences are at a level of detail not shown in the Exhibits; some of them are discussed briefly below.*

Entries are shown in this Exhibit only if the information type is not included in Exhibit B, or if the possible instruments and/or forms of response are considerably different from those shown in Exhibits A or B.

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable(A) form of responses for analysis	Suggested priority
A. Students					
1-6. See Exhibit A (A 1-6)					
7. Academic ability a. Verbal b. Quantitative c. Overall	Start and periodically during study	Interviewer or Faculty or Students	Rating scales or Self-rating scales	Single(A) or multiple(P) rating scales of each ability with normatively defined points, using upper-division students as normative base Summed rating scores(P) or Single overall rating scores(A)	Medium High
8. Knowledge base for upper-division study a. Broad areas of study b. Specialized or major field area of study and/or c. General knowledge base				Summed rating part scores(P) or Single rating score(A)	
9. Skills a. Written communication (reading and writing) b. Oral communication (speaking and listening) and/or c. General level of skills					Medium to low

EXHIBIT C (continued)

Information types	Time of collection	Information sources	Possible instruments	Preferred (P) and Acceptable(A) form of responses for analysis	Suggested priority
A. Students (continued)					
10. Satisfaction with educational experiences in lower-division study a. Features of program b. Progress toward objectives	Start and periodically during study	Students	Opinion scales	Single or multiple opinion scores based on implied scale values	Medium
11-13. See Exhibit B (A 10-12)					
B. Program procedures. See Exhibit B (B 1 & 2)					
C. Outcomes					
1-3. See Exhibit B (C 1-3)					
4. Collegiate or professional-level skills a. Written communications (reading and writing) b. Oral communications (speaking and listening)	End of study	Faculty	Ratings	Rating scores using normatively or functionally defined points	Medium
5-9. See Exhibit B (C 5-9)					
10. Assessment of knowledge in major field or area of concentration	End of study		Students or Faculty	Scores on standardized instruments yielding comparable scores (P) or Specially constructed survey tests (A) or Rating scores with normatively defined points (A)	Medium
11-12. See Exhibit B (C 10-11)					
D. Costs					
1-2. See Exhibit B (D 1-2)					

ISSUES AND OPTIONS FOR SUBSTUDY C INSTRUMENT DEVELOPMENT

This substudy can be viewed as the multiprogram evaluation version of the studies that individual programs might conduct for students who completed baccalaureate degree programs. However, the need to complete the entire substudy within the two years assumed to be available introduced so many problems and compromises that it is at best a "time sampling" study of baccalaureate students.

Any thorough longitudinal study of baccalaureate programs should extend over a minimum of four years; ideally, it should require eight or more years if the experience of parttime students is to be adequately reflected. The most important compromise assumed in the preparation of the information specifications was that Substudy C would be based only on students beginning upper-division work at the start of the evaluation study. This assumption was made for several reasons: (a) some nontraditional programs are designed only for students starting upper-division study; (b) entry to upper-division study is a major transition point for increasing numbers of students who start higher education at a community college, and this trend is likely to accelerate as higher education costs increase; (c) a number of four-year colleges award associate degrees when lower-division requirements have been satisfied, even though the students may plan to earn the baccalaureate degree; and (d) a sizable proportion of students in the defined substudy can complete the baccalaureate degree requirements during the term of the evaluation study.

A consequence of the assumption is that the only information about the lower-division experiences and programs for the students in this substudy will be what is included in college files (transcripts, etc.) or provided retrospectively by the students. A more serious limitation of the assumption is that there will be no information available for students who intended to earn baccalaureate degrees but terminated their studies while still in the lower division.

Because of these quite serious limitations, other options should be considered when the purposes and scope of an evaluation study are

more clearly defined. One would be to include an additional substudy similar to Substudy B but based on lower-division students whose degree objective at entrance is the baccalaureate. The cross-sectional approach to the evaluation of baccalaureate degree programs possible with this additional substudy would, of course, greatly increase the scope and cost of the total evaluation effort.

A second option would be to conduct a study using students who are at or near the end of their degree work when the study is started, depending entirely on retrospective information about the entire collegiate experience. Such a study would permit follow-up on the first year and a half of postcollegiate activities and progress, which could be related to their recall of the educational experience and program assessment. Other options could be enumerated through combinations of the defined substudy and options described. The cost of the various options would be markedly different.

The remainder of this discussion deals with issues that will be important for the substudy as defined, i.e., a two-year study of students starting on upper-division work.

Issues in Student Group Definition

One important issue will be that of determining the meaning of the term "upper division" for the various programs participating in the evaluation study. The distinction between upper- and lower-division undergraduates is generally quite meaningful. Usually, students earn credits at a fixed rate, and the dividing point for upper and lower divisions can be generally defined by years completed. When students pursue studies part time, perhaps encountering problems of fitting their schedules to the scheduled times of required courses, the distinction becomes less clear. For example, students may have earned sufficient credits to be in the upper division, but may not have been able to take certain courses required in the first two years. Or they may have earned all but a few of the necessary credits when they had to interrupt their studies for a time.

Colleges have often modified or waived certain requirements in response to these increasingly common problems. Some institutions, particularly nontraditional ones, have changed the ways in which they define degree requirements and recognize progress in satisfying the requirements. Years of residence is not a very meaningful criterion. In fact, for certain nontraditional programs the term "upper division" may be quite meaningless and even contrary to the programs' purposes.

It will be necessary to settle on a working definition for the identification of upper-division students which will allow for some tolerance limits in terms of time, credits, or completion of lower-division required courses. One-half of the degree requirements completed, however these requirements are stated by each program, may be the only criterion that can be used. As with so many issues, the definition may have to be negotiated with some programs to reflect unique program requirements while at the same time achieving reasonable uniformity across different programs.

A second issue in student group definition stems from the fact that students could be starting on upper-division work periodically during the term of the evaluation study. The implications of, and options for, this issue are essentially the same as those in the parallel comparable discussion following Exhibit B. Their resolution will require more complicated arrangements with the participating programs. Most colleges or programs would not routinely collect the preprogram information listed in Exhibit A at the time students start on upper-division work.

A special and major data collection effort will have to be set up for those students enrolled at the start of the study. These arrangements would then have to be continued periodically for students who enter the upper division at a later time. There may, then, be a substantial administrative and cost advantage to limiting the numbers of of "late starting" students for whom the preprogram data have to be collected. Of course, any potential saving in costs resulting from these limits would have to be weighed against the reduction in numbers of students for whom data will be available. The loss in numbers of

students will be large for certain nontraditional programs where relatively small numbers of students are at the same stage of completion of degree requirements at any given time.

Still another issue is whether the students in Substudy C who are entering the program or college for the first time should be considered as a subgroup separate from those who completed their lower-division work in the college. More specifically, a sizable proportion of students may initiate upper-division work in a program after having completed the lower-division work in another institution--a community college or at one or more four-year institutions. This proportion may vary markedly between traditional programs and nontraditional programs designed specifically for adult parttime students.

Some nontraditional programs are designed to provide only upper-division work; all the students are new to the program. Most of these programs are designed to facilitate reentry of students regardless of where and when they did their lower-division work.

This issue will go to the heart of plans for analysis of data for Substudy C. The entrance data for new and continuing students in traditional and nontraditional programs may necessarily be so different that any reasonable analysis plans will, in effect, result in separate substudies for at least four student subgroups. No speculation on whether data for the student subgroups could be pooled across programs in a single class is warranted in the absence of detailed information about the kinds and varieties of prior educational experience.

Information Content and Response Forms

As noted in Exhibit C, the program entry information types suggested for Substudy C are similar to those used for Substudies A and B. The specific instruments and resulting information would be quite different, however, simply because the students have progressed so much further in their education and have a greater variety of prior educational backgrounds.

Educational background, ability, and knowledge. More information will be needed to describe completely the preentrance educational background of students in Substudy C in ways that will permit useful grouping for data analysis. Secondary school background might be omitted entirely, but extensive information on degree-related course credits, awarding institutions, and level of achievement will be necessary.

A scheme and procedures for converting college grades and/or grade averages to a common scale will be essential. In some cases, this may require independent review and evaluation of the more unusual transcripts. The common scale may have to be one for which the points can only be qualitatively defined, or defined in very gross quantitative approximations. In any case, every possible effort should be made to arrive at a common scale. It may be, however, that no such scale can be developed because students will be found to have attended a wide array of colleges, have accumulated credits over so many years, or have been awarded credits for experiences that can in no way be considered comparable. If so, ratings of ability, knowledge, and skills at entrance listed in Exhibit C may be particularly crucial. Ratings are suggested because the use of standardized or survey tests suggested in Exhibit A to obtain this information will probably not be feasible. It may even be inappropriate in many instances. A major effort will have to be devoted to the development and tryout of these scales if they are to provide data that will be useful in the interpretation of end-of-study information.

The New York State Regents External Degree Program has accumulated valuable experience in the resolution of the kinds of problems that have been touched on only briefly here. The procedures they are employing should be studied thoroughly as a first step in planning for this aspect of the evaluation study.

Knowledge base for upper-division study. Information type A-8 is included in the specifications to provide data for the interpretation of knowledge measures that might be collected at the end of the study. The priority for such information depends, then, on the priority

attached to information about end-of-study knowledge. As indicated in the discussion of Exhibit B, such information is probably of only medium priority, so there may be less need for preprogram knowledge data than for other information shown in Exhibit C.

If such data are to be obtained, ideally one should have estimates of knowledge in both broad areas that might result from lower-division study or other kinds of learning, and in the specialized areas in which upper-division work is planned. The use of any standardized or survey tests for these purposes is quite impractical. Faculty ratings are probably the best sources of information about the knowledge base. If a choice between ratings for broad or special areas must be made, knowledge in broad areas can probably be estimated more reliably.

Communication skills. Information type A-9 has been included in Exhibit C because the methods for communication skills assessment for upper-division students will have to be substantially different from those used with students starting on lower-division work. The functional level of these skills will be important determiners of student success in upper-division work, and will also be useful in the interpretation of outcomes data. The suggested priority of "medium to low" reflects only the anticipated difficulty and cost of obtaining the information.

For those students who are continuing upper-division study in the same program or institution where they did their lower-division work, these skills can be best assessed by faculty knowledgeable about the students' work in previous courses. Such ratings will necessarily be normative for each faculty member, but it may be possible to reduce the rater variance by the use of functional descriptions to elaborate on the normative classification. An example of such a rating scale, which has not been pretested, is:

> *This student's speaking skills are:*
>
> *1. at least as good as the top 10% of the students I have had. He/she would be very effective in a graduate seminar course, and could speak well extemporaneously before a large group.*

2. *among the top 40% of the students I have had. He/she would make a good prepared presentation to a large group, and would be above average in an undergraduate seminar course.*
3. *in the middle 20% of the students I have had. He/she can respond adequately to questions in a class, and could give a satisfactory prepared presentation.*
4. *well below average. He/she can participate in discussions, but the points are not made very clearly and there is often confusion about what is being said.*
5. *in the bottom 10% of students I have had. Discussion is halting and usually quite confused; the student is seldom understood.*

Assessment of communication skills for students entering the program for the first time in the upper division will be much more difficult, since there are not useful tests of these specific skills at this level of education. It may be necessary to delay collection of this information for a month or two after entrance, so that faculty members will have time to observe the student.

Satisfaction with educational experiences. Satisfaction with lower-division program features and progress toward educational objectives (information type A-10) is included in Exhibit C for two reasons: First, such data may be of value in the interpretation of satisfaction and progress information obtained at the end of the study. Second, there may be important differences among students entering different programs on the opinion indicators that will be useful in pooling data for different programs. Exhibit C lists much less detail for these scales of satisfaction at entrance than is shown in Exhibit B for similar scales to be used at the end of the study. This is only because it seems unlikely that meaningful scales of opinions about detailed program features can be developed when students will have been enrolled in such a wide variety of programs. If time and funds permit, it would be most desirable to test this assumption before making final data collection plans.

Outcomes and costs. Although information types C and D, outcomes and costs, are essentially the same for Substudies B and C, the data collection instruments may be somewhat different. In fact, the development or assembly effort for instruments other than those used to obtain cost data might usefully include the collection of data on the extent to which the instruments differentiate between upper- and lower-division students in the same program.

The end-of-study assessment of communication skills has been included in Exhibit C because of the special problems posed if these skills are to be assessed. The skills are, of course, important to future success in graduate study and in most occupations of interest to baccalaureate degree recipients; and some decisionmakers attach importance to them as outcomes indicators. It is doubtful, however, whether improvement of these skills is a major objective for most baccalaureate degree programs, or for many students who enter such programs; there is probably little difference between the minimum skill level necessary for starting upper-division work and that necessary for postcollegiate success. Hence, the suggested priority of "medium." The comments made above about skill ratings are equally appropriate here. The definitions of functional ratings probably should be related to postcollegiate settings, however. If the students have been employed on a part- or fulltime basis while attending college, employers might be able to provide useful ratings for these students.

EXHIBIT D

Information Specifications for Substudy D
Alumni Who Received Degrees One to Three Years Prior to Start of Evaluation Study

Purposes of information collection

1. Assess outcomes, accomplishments, and satisfaction (What happens to students after they complete their programs? How has their education contributed to educational and occupational progress and satisfaction? How well satisfied are the students with what they have learned in college?)
2. Satisfaction with overall educational or college experience (What do graduates think of the program and experiences they had in college?)

Primary users of information

1. State planners and legislators, institutional presidents and deans, program administrators
2. State planners and legislators, institutional presidents and deans, program administrators

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable(A) forms of responses for analysis	Suggested priority
A. Students					
1. Sex, age, and ethnic background	During or at end of study	Alumni and College records	Questionnaire and Coding form	Frequencies in defined categories	High
2. Educational background for years preceding entrance to program					Medium
3. Educational history in program a. When enrolled, completed, time in program b. Amount credit for prior learning, % degree in program c. Area of study d. Full-/parttime study					High
4. Occupational background before and during program a. Before entering program b. While working toward degree		Alumni	Questionnaire		Medium
5. Educational objectives upon entrance to and graduation from program			Questionnaire and Rating scales	Frequencies in defined categories and Single rating scores with subjective points	Medium to low
6. Occupational objectives upon entrance to and graduation from program					
B. Program procedures					
1. Assessment of program procedures and response to needs					High
C. Outcomes					
1. Postprogram educational activities: Noncollegiate					
2. Postprogram educational activities: Upper division or graduate or professional school a. Application, admission, difficulties, transfer credits b. Area of study, change in area c. Rating of program preparation, performance d. Satisfaction with educational progress					

EXHIBIT D (continued)

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable(A) forms of responses for analysis	Suggested priority
C. Outcomes (continued)					
3. Postprogram occupational activities: Since completion of program up to current occupation a. Number b. Types, change in employer c. Expectations and reality d. Difficulties finding job, unemployment	During or at end of study	Alumni	Questionnaire and Rating scales	Frequencies in defined categories and Single rating scores with subjective points	High
4. Postprogram occupation: Current a. Title, employer, hours, income b. Closeness to goals c. Relation to area of study d. Rating of job satisfaction, work values, progress e. Rating of program preparation, performance, circumstances as direct result of program		Alumni and Employer			
5. Ultimate educational and occupational objectives		Alumni			
6. Gain in abilities, skills, knowledge as result of program					

ISSUES AND OPTIONS FOR SUBSTUDY D, INSTRUMENT DEVELOPMENT

For some decisionmakers, this may be the most important of the substudies because it is directed to assessment of outcomes as reflected after graduation. The purposes of the substudy can be easily oversimplified if insufficient attention is given to the differences among the various types of degree programs with respect to purposes and features. Ideally, the study should be done as a longitudinal extension of Substudies B and C, but such long-term studies do not seem practical as a means to provision of information for decisionmaking. The shorter cross-sectional study suggested here could easily be done within the two-year time limit.

Completing the alumni substudy within the two-year time period will require that information about, and assessment of, program related experiences be dependent on recollection. Many alumni will have had some of these experiences as much as ten or more years prior to the start of the study. Opinions about satisfaction will relate to complex and detailed program features that may have extended over many years. Dependence on these recollections may again reduce appreciably the value of the information for an understanding of differences between traditional and nontraditional programs.

Data Collection Cycles and Alumni Subgroup Identification

In designing this substudy, a choice between two data collection schedules is possible. The first possibility is to collect all the information listed in Exhibit D only once, analyze and interpret the data, and complete the report. This could be done any time during the two-year term of the evaluation study, although it would be preferable to do it rather early to minimize the time span for which alumni recollection would be a factor. The alternative data collection strategy is to carry out this data collection task near the start of the study, then to update information types C-2, 3, and 4 about a year-and-a-half later through a second data collection cycle.

The relative simplicity and limited costs are, of course, the principal arguments for the first of these two possibilities. It could conceivably be done within a six- to nine-month period if one were to forego the potential usefulness of data from the other substudies in the interpretation of the results of Substudy D. More important, perhaps, given the magnitude of the initial data collection tasks in Substudies A through C, there are great administrative advantages in delaying information collection for the fourth substudy until, say, the seventh or eighth month of the evaluation study.

The arguments in favor of the second alternative rest on the value of getting a more complete picture of what has happened to the alumni. Relationships between the collegiate experiences and the documented activities and experiences during the additional time could be examined. Although the second possibility would be more expensive than the first, it would capitalize on the initial cost of locating and contacting the alumni. Arrangements for the second data collection cycle could be appreciably simpler than for the first, since the alumni could be encouraged to keep the evaluation study staff informed of changes in jobs and location after the first contact with alumni has been made.

The substudy is limited to recent alumni (those awarded degrees one to three years prior to the start of the evaluation study) because a number of nontraditional programs have been awarding degrees for only a few years; it will also be easier to locate and contact recent alumni. The limitation may also maximize identifiable program effects since there will be less variance in educational and occupational progress attributable to postprogram opportunities and accomplishments.

The choice of degree programs and students to be included in Substudies A, B, and C will automatically determine four different groups of alumni: those who received associate degrees and did not continue with formal education; those who received such degrees and continued formal education; those who received baccalaureate degrees and then terminated their formal education; and those who continued formal education after receiving the baccalaureate degree. It is

likely, however, that traditional and nontraditional programs will have very different numbers of alumni in each of these four groups. The potential cost of locating, contacting, and following up alumni will dictate some kind of stratified random sampling for most traditional and many nontraditional programs. The size of the groups for some of the nontraditional programs will require that all of the alumni be located and contacted.

The definition of strata that would be useful in sampling is not at all clear. The choices of strata will be limited by the kinds of information colleges have about their alumni, which will vary from one program to another. It is possible that the only strata that can be used are elapsed time since award of degree, age, sex, and ethnic groups; these may have little relevance to the variables of concern in the evaluation study. If most programs have information about purposes for seeking a degree, gathered at the time of entrance or almost any other time, then a scheme for classifying such data in terms that can be applied to all programs should certainly be developed for use in sample selection.

The Bureau of Social Science Research (BSSR) is currently conducting a study similar in many respects to Substudy D, surveying alumni of nontraditional programs. In view of the investment already made in that study by the National Institute of Education (NIE) and of the potentially high cost of Substudy D, it will probably be most efficient to concentrate attention and available funds on the other substudies outlined in this report.

Information not provided by the BSSR study would be obtained from Substudy D, for example, employer ratings of program graduates. But the BSSR staff is gaining experience and knowledge about the conduct of a multiprogram alumni follow-up study that will be invaluable in the planning of Substudy D, if it is to be done. At the conclusion of their study, they will probably have recommendations for additional questions that need answers in a subsequent study. At the very least, initiation of Substudy D should be delayed until the data collection and analysis phase of the BSSR study has been completed.

Information Content and Response Forms

The information types listed in Exhibit D are for the most part self-explanatory. Nearly all the types listed in Category A--"Students"--will have to be provided retrospectively by the alumni, as noted above. Separate questionnaires probably should be used for associate and baccalaureate degree recipients, with branching in each for terminal and continuing education alumni.

The important issues to be resolved in planning for information collection will revolve around the priorities assigned to the information types, the amount of respondent time that can be expected, the amount of detail to be collected for each information type, and the extent and nature of the follow-up on nonrespondents.

Most colleges have conducted alumni surveys on several occasions. Their questionnaires range from brief one-page forms asking essentially for job titles and additional degrees earned since graduation, to multipage forms covering college experiences, employment, education, hobbies, political orientation, and the like. For a multiprogram evaluation study, some reasonable compromise will have to be reached. If questionnaires are the principle instrument to be used, a half-hour of respondent time may be about optimal. If telephone interviews are to be used as a procedure for follow-up of nonrespondents, then an even more abbreviated schedule will be required for these alumni.

Virtually all the information listed in Exhibit D is important to a meaningful interpretation of data on subsequent progress and satisfaction of alumni as they are related to educational experience in degree programs. The suggested priorities of "medium" or "high" are related to the amounts of detail that might be obtained, rather than to the need for the information itself. Even within high-priority information types, potentially useful data detail may have to be sacrificed in the interests of maximizing the rate of response by limiting responded time required to complete the questionnaire. It would be highly desirable to develop alumni opinion and rating scales that have useful measurement characteristics similar to those discussed in association

with Exhibit A. If they could be developed, however, they might require so much time that other useful data could not be obtained. Therefore, it may be necessary to use much simpler scales for the alumni survey.

No information relevant to the cost of program participation, or of the sources for meeting these costs, is included in this substudy. More current information would be obtained in Substudies B and C, and it seems likely that the estimates obtained in those substudies would be more reliable.

Program staff may have different priorities for information collection than those suggested. The final decisions will have to be negotiated with the program/institution staffs, since the latter will have both an important stake in the results of the study, and a crucial role in the success of the data collection effort.

EXHIBIT E

Information Specifications for Substudy E
Professional Perceptions of Programs Reported by Faculty and Administrators

Purposes of information collection

1. Acceptance of degrees and credits (Will other institutions accept at face value the degrees and credits awarded by this program?)
2. Assessments of intellectual content, quality and rigor of program instruction, assessment and other procedures (How credible is the program to professional personnel within and outside the program?)
3. Faculty and staff satisfaction with personal and professional development (Are there special faculty requirements or incentives?)
4. Impact (What impact have these programs had on higher education practices?)

Primary users of information

1. State planners and legislators, institutional presidents and deans, program administrators, federal R & D program staff
2. State planners, institutional presidents and deans, program administrators
3. Program administrators and faculty, higher education specialists
4. Federal R & D program staff, higher education specialists, state planners, program administrators

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) form of responses for analysis	Suggested priority
A. Acceptability of degrees and credits					
1. History and current status of accreditation	As appropriate for each program	College & program staff	Interview schedule and Accreditation committee reports	Interview summaries Excerpts and summaries	High
2. Approvals by institutional senates, committees, and officers			Minutes and records	Summaries and excerpts	Medium
3. Acceptance of associate degrees by four-year colleges*			Interview schedule and Documents	Interview summaries or Frequencies in defined categories	High
4. Acceptance of baccalaureate degrees by graduate schools*					
5. Acceptance of degrees by employers for hiring and promotion*					
B. Faculty and administrators perceptions					
1. Student ability, preparation, interest, and performance	End of each year of study or end of study	Program and nonprogram faculty & administrators	Rating scales	Multiple and summed rating scales using normatively defined values	Medium
2. Program faculty qualifications, teaching and research interests, and standards for student performance			Rating scales	Multiple rating scales using subjectively defined end and intermediate points	
3. Program procedures for recruitment, admissions, instruction, and uses of resources			and Interviews	and Interview summaries	
4. Total program strengths and weaknesses					

* Some or all of this information may be provided more directly by alumni as part of Substudy D.

EXHIBIT E (continued)

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) form of responses for analysis	Suggested priority
B. Faculty and administrators perceptions (continued)					
5. Extent of knowledge about program features or basis for ratings	End of each year of study or end of study	Nonprogram faculty & administrators	Questionnaire and Interviews	Frequencies in defined categories and Interview summaries	Medium
6. Personal and professional rewards and penalties for program participation	↓	Program faculty	Interviews	Interview summaries and excerpts and/or frequencies by defined categories	High
C. Description of program faculty and administrator qualifications	↓	Program staff	Interviews and/or Coding forms	Interview summaries or excerpts and/or Frequencies in defined categories	Medium
1. Educational background and teaching experience 2. Research experience 3. Administrative experience 4. Other relevant experience	↓	↓	↓	↓	↓
D. Program impact	↓	↓	↓	↓	↓
1. Modification or adaptation of procedures by other programs	↓	Program and non-program staffs	Interviews	Interview summaries and excerpts	Medium to low
2. Grants for special studies or dissemination of program procedures	↓	↓	↓	↓	↓
3. Publications in professional journals or books describing program practices and accomplishments	↓	↓	↓	Bibliographical material	↓

ISSUES AND OPTIONS FOR SUBSTUDY E DATA COLLECTION

This substudy is directed to several quite different program evaluation issues: acceptability of degrees, collegial perceptions of the program, roles and qualification of program staff, and program impact. Higher education traditionally has depended on peer judgments of professional qualifications, reputation, and productivity as a major--perhaps the major--indicator of program quality. Thus, this substudy may provide the most important evaluation information in terms of the priorities of the higher education professional community. This community is too broad, and is shifting too constantly over time, to be defined as a decisionmaking group, but there can be no question that it has a very important, if nebulous, effect on decisions.

The outcome of this substudy will be the organization of evaluative judgments about programs designed to have effects on other people. Thus, it is substantially different from Substudies A-D, which were defined in terms of groups of students who are to be, or have been, affected by the programs.

In the earlier substudies, those providing the information are presenting perceptions of programs designed for their benefit, and reporting on the effects of the programs. In this substudy, those providing the information are rendering judgments about the programs they may have designed or at least shaped (program staff), or they are expressing opinions about programs which in some sense may be competitive with ones in which they are active participants (nonprogram staff). In short, those people providing information for Substudy E will have interests in the programs that are quite different from those that influence the information given by students or alumni.

In a sense, Substudy E is intended to collect and relate the judgments of three groups of professional personnel who can view the programs from different perspectives: program faculty and administrators, nonprogram faculty and administrators from the same college or institution, and an independent group of professionals who were asked to evaluate the programs for an accrediting agency. The varieties of

organizational structures in higher education and the nature of collegial review and evaluation are such that these distinctions in perceptions often will not be very clear when arrangements are being made for information collection at the several programs. Insofar as possible, however, the need to make them should be kept in mind as the information is collected, analyzed, and reported.

Identification of faculty and administrators from whom information is to be solicited will be an important issue to be resolved as arrangements are made for data collection. The selection process may be much more important in determining the results of Substudy E than was true for respondent selection in the other substudies. In many instances, the choices will be fairly evident once the structure of a program is known, particularly for the program administrators. The choice of faculty members will not be so obvious, particularly for programs using adjunct faculty, unless it is possible to contact all faculty members.

If information about traditional and nontraditional programs in the same institution is being collected, it may be useful to obtain ratings of both programs from the same faculty members. Faculty members in a nontraditional program could be asked to rate this and the "comparable" traditional program, and vice versa. Such a procedure will be more feasible if campus-based programs are offered in off-campus locations than for other nontraditional programs. There should be faculty and administrators who have first-hand knowledge of both. Whether or not such arrangements can be made, however, an important consideration in the interpretation of faculty ratings of programs will be the extent of knowledge the rater has of a program in which he or she is not an active participant. If it were possible to get an indication of knowledge of program features, this could be an important consideration in the choice of those to be asked to provide ratings of various program features.

The information types listed in Exhibit E are essentially self-explanatory. The actual instruments to be developed may have to be tailored to each program so that terminology is meaningful. The possible analysis of data will be so dependent on the actual choice of respondents that little more can be offered at this time.

EXHIBIT F

Information Specifications for Substudy F
Program Costs, Sources of Funding, and Financial Viability

Purposes of information collection

1. Historical and/or current program cost information (How much do the programs cost? What are the one-time development costs of such programs?)
2. Financial viability of programs (What have been the sources of funding for the development of the program? What are the sources of funds for current operations? Will the programs be able to continue operations without special federal, state, or foundation support?)

Primary users of information

1. State and federal legislators and planners, federal research and development agencies, institutional presidents and deans, program administrators
2. Same

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of responses for analysis	Suggested priority
1. Current program costs for geographically separate units (if appropriate) and for total program	End of each year of study or End of study	Business officers, Program directors, Directors of separate cost centers	(See discussion)		High
a. Instructional					
b. Administrative					
c. Facilities					
d. Overhead					
e. Total					Medium
2. Student/staff ratios for program and/or for geographically separate units					
a. Student/faculty					
b. Student/administrative staff					
c. Student/adjunct faculty					High
3. Program charges to students					
a. Tuition					
b. Fees					
c. Other					
4. Financial aid provided to students		Financial aid office	Questionnaire	Frequencies in defined categories	Medium
a. Federal loans, grants					
b. Work-study programs					
c. State scholarships and loans					
d. Program or institutional scholarships, grants, loans, etc.					
e. Other sources					
5. Income by source		Program director and Business officer and Program documents	Interviews and Coding forms	Interview summaries or extracts and Document extracts	High
a. Tuition and fees					
b. Local and state taxes					
c. Special federal and state support					
d. Foundation and other private support					

EXHIBIT F (continued)

Information types	Time of collection	Information sources	Possible instruments	Preferred (P)and Acceptable (A) forms of responses for analysis	Suggested priority
6. History of program costs and sources of funding for program planning, facilities development, operation	As convenient	Program director and Business officer and Financial reports	Interviews and Coding forms	Interview summaries or extracts and Document extracts	Medium
7. Estimates of future income from various sources a. Short-term b. Long-term		Program director or Various institutional administrators	Interviews	Interview summaries or extracts	High

ISSUES AND OPTIONS FOR SUBSTUDY F

For some decisionmakers, this substudy will be the one of greatest interest. Nearly all decisionmakers will have some interest in it. Frequently, cost data may be the only information used in arriving at a decision about an educational program; the decision is based on educational values and philosophy, supplemented by cost data in a very simplified form. And the viability of long-term funding of programs may determine whether other decisions are necessary.

Program Costs

Two approaches to the collection of cost data were considered for Substudy F. One approach would be to collect quite detailed data on the costs of various separate features as well as the total costs for each of the programs included in the study. The second approach would be to obtain cost data only for total programs and for geographically separate units, essentially ignoring the differences in the nature and combinations of features used.

The first of these approaches is particularly appropriate for program management decisions necessary if program costs must be reduced. Analyses of the cost data for separate features in a multiprogram study could also be helpful to program administrators by identifying potential efficiencies in program procedures that might merit further exploration for potential adoption. There is a real risk, however, that so much data would be generated that a comprehensive interpretation would be virtually impossible. Additionally, interpretation of the cost data would require discussion of program features and costs out of the context in which they occurred in the evaluation study so that conclusions could be seriously misleading. Until much more is known about the importance and effects of various program features and costs on a program by program basis, such detailed cost studies would probably best be done by the staffs of each program.

The information specifications for Substudy F are based on the second of the two approaches--that of collection and analysis of cost

data only for the total program. These total cost data should be presented and analyzed for each geographically separate unit if the cost accounting procedures employed by the sponsoring institutions treat each unit as if it were a separate program. For interpretation of the cost data, it may also be useful to break down the total costs by major categories of objective of expenditure, as shown in Exhibit F.

To obtain an overview of program costs, important variations in the definition and computation of costs, and in the assignment of costs to various categories, have to be overlooked. Such variations may be of particular importance in studies of nontraditional programs, and can be ignored only at the risk of possible misinterpretation of the cost data. State and institutional accounting practices are at different stages in responding to the accounting problems posed by the introduction of new structures, procedures and requirements of some nontraditional programs.

The problems and difficulties of doing comparative cost studies in higher education are well known. They are not unique to nontraditional education programs, but such programs frequently represent extreme instances of the problems. The problems are discussed in detail in Medsker et al. (1975), and the authors concluded that only the most general statements about costs of nontraditional programs could be made. The interpretation of comparative cost data for both traditional and nontraditional programs would require even more caution than was advocated.

To indicate the magnitude of the problems that will be encountered in collecting and analyzing data for information type 1 in Substudy F some of the most important problem areas are listed and discussed below.

Full-time-equivalent units across programs. If most courses in a program or institution meet for a fixed number of hours each week, it is possible to derive an approximation (credit hours) to the amount of weekly effort required for students to complete the course, or an approximation of the number of "class contact" hours between students

and faculty. Using one of these or other approximations, which are kinds of "average time per student per course," it is possible to arrive at an "average course load" if a student is engaged in fulltime study for some calendar period of time, say a quarter or semester. This course load then becomes the definition of a fulltime student, which can then be used as the basis for "unit cost" analysis. The course load of students doing parttime work can be converted to a percentage of the fulltime load.

The uses of these "averages" are based on a number of assumptions. which vary from one institution to another and even among divisions of a single complex institution. Still, in the absence of any more rigorous procedures, the averages have been found to be serviceable for administrative purposes in most programs and institutions. The rationale and philosophy of most programmatic innovations require freedom from the constraints imposed by fixed classroom hours and group instruction; the assumptions underlying the use of class hours to approximate student effort simply do not apply, especially for the programs involving self-directed study or individualized study.

If students meet with faculty only occasionally for blocks of time, or for periodic individual study planning conferences, the initial building block for estimating total student effort is missing. The faculty contact hours represent a very much smaller proportion of the total student effort, and learning time is largely a function of individual student interest and motivation outside the program or institution walls. No meaningful assumptions can be made about the total time a student can or will devote to learning in a calendar period. In fact, for many of these programs the calendar period--quarter, semester, or whatever--is not even meaningful because the student has flexibility in setting completion dates.

Nontraditional programs have adopted a variety of strategies to replace credit hours. Some can rather easily adopt at least the form of campus-based conventions, relying on individual faculty members or committees to convert assumed student effort into the credit units

employed on the campus. The programs that stress individualization often have simply discarded the established credit-hour currencies and have replaced them with a completely new system for recording effort and achievement. They generally have adopted some entirely arbitrary conventions that relate their new system to other established credit-hour systems, in order to provide their students and alumni with the common currency needed for graduate school admission and occupational advancement.

However the units are derived to define fulltime student effort and achievement, the assumptions on which they are based should be reviewed carefully for each program and institution to assess their implications for the use of the fulltime equivalent as if it were comparable for all programs.

Faculty age, length of service, salary scales, and participation in innovative programs. Faculty salaries are by far the largest single cost element in educational programs. Allowance must be made for factors affecting these costs in ways that could be biased for some programs if the results of a comparative cost study are to be interpreted correctly and used in meaningful ways.

Innovative programs are often staffed with a high proportion of younger faculty members for a variety of reasons. Academic salaries are closely related to age and length of service, with a ceiling at each rank that may be reached only after a number of years of service. Thus, faculty costs for new programs may be relatively low in the early years of the program, due partly to the use of younger faculty. They may, however, increase rapidly for several years before salary ceilings are reached even if the program is not enlarged. As the life of the program lengthens, faculty will begin to qualify for higher ranks and possibly tenured positions, with resulting salary increases; but a good many years may be required before the distribution of faculty ranks in the new program is the same as for older programs. Ultimately, successful innovative programs will have to bear the costs of a full range of faculty ranks. Thus, faculty costs may be lower for new or nontradi-

tional programs, due to the design of the program, but also as an artifact of the age of the program.

Faculty teaching versus research time. In most four-year institutions, teaching loads are assigned recognizing that research is an important part of the faculty member's job. The overall proportion of faculty time (and salary) allocated to research instead of teaching will vary across institutions according to the mission of the institution or college. Within institutions, it will also vary according to faculty rank, with the effect that the higher the rank and salary, the larger the proportion of time that will be allowed for research. A number of institutions include total faculty salaries in the cost of instruction, because research time is viewed as contributing to the teaching effectiveness of the faculty members, and because in many instances there may be no other way of recovering the costs of faculty research time. If this is the practice of an institution, it will be necessary to unravel the research and teaching costs incurred in faculty salaries in order to achieve some comparability across programs.

Programmatic instructional innovations are less likely to be undertaken at institutions where research is an important aspect of the institutional mission. Additionally, faculty who have lower teaching loads in order that they may pursue research interests are less likely to have the time required for participation in innovative programs. The calculated instructional costs of traditional programs may, therefore, include a larger proportion of faculty research time than is the case for nontraditional programs. Eventually, faculty in innovative programs should and will want some reduction in teaching loads so that they too can devote more time to research interests. In the early years of nontraditional programs, when interest in comparative cost data is highest, detailed information should be gathered for each program on the methods for accounting for research time and relative proportions of faculty time devoted to research and teaching.

Faculty overload, adjunct faculty, and tutorial staff. Innovative programs may have options for faculty employment not available to established or traditional programs. Some of these options will serve to reduce program costs in the first years of the new programs, but will be only temporary; others may represent more permanent cost reduction strategies. In any case, the nature and long-term viability of the options must be considered as comparative cost data are collected and analyzed.

Most innovative programs necessarily have to start out as relatively small scale efforts; frequently they are the outgrowth of interests of faculty members who already have fulltime teaching and research positions. A staff with the breadth of backgrounds, interests, and knowledge necessary for a balanced program may have to be assembled quickly. To accomplish this, faculty are often recruited to assume responsibilities in the new program in addition to their fulltime loads. Payment for this additional teaching load does not have to include benefits and may not carry with it allocations for various indirect costs which are fully absorbed by the regular salary payments. Thus, the costs of these "overload" services may be less than the "hourly" (or per student) rate that is paid for the onload services of the faculty when teaching in traditional programs. In effect, the innovative program is "subsidized" by the faculty or sponsoring institution.

Many nontraditional programs also are designed to use adjunct faculty whose principal income is derived from sources outside the program. In effect, these people also provide services to the program on an "overload" basis. In some cases, this adjunct faculty may volunteer their time because they see the service as an important part of their regular work, or they may provide their services at a cost lower than the rate of pay for their other occupation. If the arrangements for these or similar services (i.e., the same kinds of services but not necessarily the same individuals) can be considered to be permanent, then the costs will be unchanged. If, however, volunteers eventually have to be replaced by paid adjunct faculty, or if needs being met by adjunct faculty

increase so that employment of fulltime faculty is required as the numbers of students increase, then the costs of these services in the early years will be underestimates of future costs.

Finally, it is often both desirable and possible to provide individualized instruction, particularly in skills development, by employing qualified but lower paid tutors rather than professionally trained faculty. While this same option may be available for traditional programs, they may not be as free to pursue it because of commitments made to professional faculty. The employment of tutors may be dependent on funds that have to be used for faculty salaries.

Telephone, mail, travel and media costs as instructional costs. Innovative programs frequently use faculty in new roles that require closer faculty/student relationships. These roles reduce the time required for campus contact with groups of students, but make increased use of communication channels not traditionally associated with instruction. If the use of these channels entails large costs, it will be necessary to include these costs as part of the instructional costs for nontraditional programs.

Individualized instruction will often require less faculty time for face-to-face meetings with students because the student has a larger responsibility for the learning that occurs. However, periodic lengthy telephone conversations or correspondence may be required in lieu of these on-campus contacts. As the student progresses in the planned study, extensive review, advice, and redirection may be useful, but distance from the campus may preclude face-to-face contacts. The time required of the faculty member for this extended consultation will be great, of course, but it is the kind of time for which it is difficult to assess charges. Further, however, there will be disproportionally higher costs for mail, telephone services, and secretarial support which are integral parts of the costs of instruction.

Other nontraditional programs may rely heavily on the use of various kinds of media presentations (tape cassettes, film strips, etc.) in lieu of classroom lectures and discussion. These materials

are not used to supplement instruction, as if often the case in traditional programs, but to replace classroom instruction. Thus the costs should be included as direct charges for instruction, but some account should be taken of the fact that supplemental media services provided for traditional programs are not required for some nontraditional ones.

Finally, extended degree programs may require frequent faculty travel to geographically separate locations. Such travel is necessary for traditional programs, of course; but for these programs the students do the traveling and pay the costs. Faculty and administrative costs for travel have to be charged to the programs, but any cost analysis interpretation should consider the fact that the total cost of instruction for a group of students may be reduced appreciably by the great reduction in private travel costs.

Cost classification for specialized program services. If innovative programs are designed in part to attract students not reached by traditional programs, specialized noninstructional services are often required. Frequently, these services are provided by traditional programs as well, but the services must be delivered in different ways by different kinds of personnel if they are to be effective for nontraditional programs. Programs and institutions differ in the classification or assignment of costs for these services, and often they are not included as instructional program costs.

As an illustration, all colleges have an active publicity and student recruitng service, which frequently involve travel, student interviewing, and presentations to off-campus groups. The costs of such services may be generated by public information, admission or other offices, and assigned to administrative services or institutional overhead. They may be partially or fully recovered through application or registration fees, or from public funds provided for institutional administration. If so, they are not usually included in the costs of instructional programs offered in a department or school within the college or university. An innovative program may require essentially the same services, but they must be directed to potential

students not reached by the ongoing institutional services. New publicity materials specific to the program and clientele are needed, and additional specially selected and trained staff are required. The costs for these new materials and staff will not be offset by any reduction in costs for services being provided by the college, and may be very appropriately charged to the new program. Any comparative cost analysis should, however, make allowance if established programs receive similar kinds of services "free."

As another illustration, admissions and/or registrars offices usually spend sizable amounts of time and money collecting, reviewing and processing transcripts, letters of recommendation, admission test scores, and so on for new students. The costs for these administrative support services are seldom assigned to instructional programs or recovered through tuition charges. Some innovative programs require similar kinds of information but the information is used as an integral part of the instructional process. Transcripts and other information provided by entering students are reviewed and evaluated as part of instructional program planning. Faculty time may be devoted to assisting the students in the assembly and preparation of summaries of prior learning. In short, the entire entry information collection and processing are part of the instructional program, and costs should be assigned to the program.

One more example may be helpful in illustrating the ramifications of these issues for comparative cost studies. Students who are starting or resuming higher education after a long lapse often want some special orientation or assistanc to get "back into the swing" of academic work. Others who are undertaking degree work as a prelude to starting a new career or changing occupations may want advice and counseling about career requirements and choice. In some highly innovative programs, all entering students may need extensive orientation by faculty and staff to the requirements and procedures of the program, which will be markedly different from those with which the student is familiar from past experience. Orientation and a variety of counseling services are,

of course, provided by all colleges for all students. The orientation is often quite brief, however, and frequently is the responsibility of administrative staff and students. Counseling services may be available in many different places--student health services, placement offices, psychology department clinics, etc.--but are seldom located in academic departments or programs. The assignment and recovery of costs for the services is equally diffuse. The provision of orientation and counseling services in the context of the new program will generally not result in a reduction in the need for such services provided by a college or university. In other words, the new services will generate "add on" costs attributable directly to the creation of the new program. A comparative cost study must recognize that similar services are provided to all established instructional programs, but the cost accounting system may not charge the programs for them.

Space and equipment costs. When new programs are offered in off-campus locations, the allocation of costs for space and equipment usage, which is never a simple matter, can create a number of complications for the interpretation of comparative cost data. Space for such programs may be rented or obtained from other institutions or agencies (e.g., high school buildings, community center buildings, libraries, or other colleges) with or without rental charges. Seldom, if ever, is any capital construction or purchase involved; whatever the costs, they would generally be treated entirely as current operating expense.

On-campus programs, on the other hand, most often make use of capital buildings and equipment. The constructions costs of some of these buildings may have long since been fully amortized, so the only current operating charges are for building maintenance and operation. The costs of other buildings may include interest and bond retirement charges in addition to maintenance and operation. Sometimes the capital costs are obligations of the state met from general revenue funds and may never be reflected in current operating expenses. The most common practice is probably that of aggregating all debt service,

maintenance and operating expenses for a campus and prorating them to departments or programs on the basis of total costs of one kind or another. In any event, the charges to an on-campus program may be appreciably different from the charges if similar space had to be rented off-campus.

There is such variation in the kinds of space charges incurred by different programs, and the accounting procedures used to assign these costs to programs, that it may not be sufficient simply to note that these variations exist when interpreting comparative cost data. Some standard rules for adjustment of the charges may be required to achieve any approximate kind of comparability.

Cost accounting and costing models. This partial listing of problems and issues that must be dealt with in comparative cost studies can only suggest the magnitude of the problem. The list could be expanded almost without end if all the possible variations in cost accounting practices for different programs and institutions were considered. There are very good reasons why comparative cost studies could not be located.

It seems most likely that no study is possible if one were to use actual cost data for each of the programs to be included in the study. There are a number of cost models for higher education developed to apply common methods of classification and analyses using best estimates for various kinds of costs as provided by colleges or programs. The use of such models as the basis for a comparative cost study may be more fruitful than actual cost comparisons because the errors attributable to variations in accounting practices may tend to average out over the large number of cost items. Clearly, a great deal of experience would be required before any conclusions about the utility of costing models could be drawn.

Student/Staff Ratios

Information type 2 is included in Exhibit F as a second, much cruder indicator of costs. Personnel costs may account for at least 70 percent of the total program costs. These ratios may therefore be useful as

approximations to the *relative* costs of programs, since they are free of many of the problems and program-to-program variations in accounting procedures discussed above.

Other Information

Information types 3 and 4 could be useful in the interpretation of data on private costs to students, and possibly in the description and understanding of various program differences.

Information types 6 and 7 are closely related, and seem essential if long-term financial viability of programs is of interest.

IV. CONCLUSIONS AND RECOMMENDATIONS

We have presented preliminary specifications for the information necessary to meet the high priority needs on which there appears to be some consensus of decisionmakers. In preparing these specifications, we have given attention to the feasibility of collecting such information, and have limited them to the minimum information necessary for adequate interpretation of data that might be collected. This required numerous compromises which would well be debated. It seems clear that the total study implied in this report would be too large in scope to be undertaken in the reasonably near future.

During the preparation of these specifications, we reached a number of conclusions about approaches to the important task of independent or external evaluation of traditional and nontraditional degree programs. Some of our judgments based on previous work were reenforced. We are limiting ourselves to a few crucial recommendations, recognizing in every case that there are alternative courses of action that could be presented forcefully.

1. A separate and major effort in instrument development should be completed at least through a pretest stage before any plans for a multiprogram evaluation study are made.

The need for instrument development in advance of detailed study planning is particularly important in the context of an evaluation study of the magnitude of the one presently being considered. The traditional approach to evaluation instrument development cannot be justified. The cost of the study will be too great to simply incorporate instrument development as a task in a total study plan.

2. The instrument development effort should not be initiated without a prior systematic review of the progress and schedule for relevant projects now being carried out independently.

We have become aware of a number of projects now going on that have great potential value for a multiprogram evaluation study. Some may result in usable instruments in one or two years; others may require more time. Particularly relevant development work is being done by the American College Testing Program, Bureau of Social Science Research, Educational Testing Service, Empire State College, and the National Center for Higher Education Management Systems. The projects are being variously financed by the Fund for the Improvement of Postsecondary Education, the National Institute of Education, and nongovernmental sources. A number of colleges and universities are also embarked on similar efforts, using their own funds to solve their own problems. These program-specific efforts may provide at least prototypes of instruments that could be useful as a base for the development of instruments for more general use in an external study.

Coordination and even some joint planning of these efforts with attention to the external evaluation of programs using common instruments is greatly needed. The task will be so complex that the creation of a consortium of public and private institutions and funding agencies might be required both to coordinate ongoing efforts and to identify important instrumentation needs not now being addressed.

3. There is a great need for well developed and tested devices to assess postcollegiate occupational and educational progress in relation to the collegiate experience and outcomes.

Most of the effort to develop better indicators of the quality and results of collegiate experience has understandably been directed to the assessment of knowledge and skills. The recent emphasis on competency assessment has renewed and redirected these efforts, but they are still focused on cognitive and behaviorally definable short-term outcomes. At the same time, many decisionmakers now view these

outcomes only as means to more important ends of personal and career success and satisfaction. These ends are, however, stated in only the most general terms, and the assessment of student status or progress with respect to these goals only reflects the imprecise definition of them. The most frequently used approach is to provide a limited number of such statements, and ask people to indicate status by choosing a descriptive word or phrase that can be defined only subjectively by the respondent. Such forms of response are limited in their usefulness for analysis, and generally do not provide sufficient discrimination among people or programs to support any conclusions about the effectiveness of different programs in meeting the needs of clearly different groups of students. It is necessary to develop opinion and rating scales that provide more definitive statements of ends and less subjectively defined forms of response.

This recommendation should not be taken to imply that the assessment of cognitive outcomes of collegiate experience are not of crucial importance. Rather, a parallel effort to define and measure more accurately some of the longer-term outcomes is needed. The effect of college attendance on success and satisfaction in both personal and societal terms should be a primary concern to decisionmakers at all levels. The question is too important to be dealt with only casually in evaluation studies. Well developed and tested opinion-and-rating scales are sorely needed.

> 4. Survey or broad-band instruments for the assessment of cognitive or competency outcomes should be developed to reduce the dollar cost and required student time for program evaluation.

Historically, much of the test development in higher education has been designed to aid in student selection, diagnosis, and placement. This has necessarily required the use of relatively long and costly instruments that will yield scores reliable enough for decisions about individuals. Program assessment does not require such extensive information about each individual. The use of shorter tests, perhaps

combined with item sampling techniques, would in many cases be entirely sufficient.

> 5. Rules or criteria for the determination of comparable programs should be developed prior to the design of any evaluation study. Such rules should encompass not only comparison between traditional and nontraditional programs, but also the comparison among various programs in each of these categories.

During a period of declining resources for higher education, decisionmakers must choose to fund some programs and terminate others, regardless of comparability. This dilemma is implicit in decision-oriented evaluation. It can be resolved only if boundaries can be established within which different programs will be considered to be comparable for the interpretation of information not specifically dependent on particular program features. Tolerable boundaries for acceptable comparability should be established early in the study planning phase as a guide for the detailed data collection and analysis planning.

Programs are designed by institutions to provide services consistent with the institution's purposes and needs. These programs are used by students in ways that will meet student needs. The rules for determining comparability of programs, then, could be based on the programs as designed and described by the institutions, or as they are described by the students for whom they are provided, or on the combination of student and institutional purposes and plans. The last two of these three require information that could be obtained from an evaluation study, but preferably should be obtained prior to the formal study to avoid confounding the application of the rules and the interpretation of the data used for comparisons.

> 6. Rules or criteria to determine the inclusion or exclusion of innovative programs in a multiprogram evaluation study should be established.

One of the greatest threats to innovation is the premature evaluation of their effectiveness. Any innovation that is worth trying

requires time for trial-and-error learning. Innovative programs should not be expected to have achieved their major goals and objectives in a short period of time, even though funding proposals and requirements may force the staff to make high-sounding promises. Certainly, if programs are to have a relatively long life they should not be evaluated (except for formative evaluation) until the novelty of the program and the sense of exploration and commitment of the original staff are themselves no longer important program features. Additionally, any useful summative evaluation study should attend to some of the postcollegiate effects for students enrolled after the program has been stabilized. If alumni are to be included in a multiprogram evaluation study, only programs that have graduated at least five "classes" should be considered.

Program size and the acceptability of the important features within the boundaries of comparability discussed above would also be factors to be considered in establishing these rules.

> 7. Any plan for a large-scale multiprogram evaluation study should be pretested on a small scale before the total plan is put in final form and arrangements for implementation of the plan are made.

An evaluation study of the scope implied in our information specifications will be too difficult and expensive to undertake without a small-scale empirical tryout of the plan for the study. However carefully the planning may be done, the problems that will be encountered will be too numerous and complex to be anticipated and thought through to successful resolution. The process of experiential learning by evaluators about the details of unanticipated problems and acceptable solutions should not be part of the evaluation of the programs. The importance of the study to the programs and the total cost of the evaluation study require that a pilot study be done in advance.

If this recommendation is accepted, we suggest that no evaluation study more complex than a combination of Substudies A and B (associate degree students, including dropouts and transfer students) be undertaken for a pilot test. We would suggest further that such substudies

include no more than three nontraditional and three traditional programs. Such a limited study could provide some data for tentative and limited conclusions about the effects and uses of these programs and others quite similar to them. They would be more than just a number of institutional or program self-studies. Their primary value, however, would lie in the experience gained in the use of the instruments and in the problems and issues to be resolved in making arrangements with the institutions and understanding the data interpretation problems when data are collected at a variety of programs.

8. Initiatives for comparative cost studies, and efforts to develop cost effectiveness indicators for one or more programs should be separated from any multiprogram outcomes studies.

The present lack of uniformity in cost accounting practices employed in higher education reflects the institutional diversity and autonomy that characterize higher education in the United States. Colleges and universities supply certain kinds of cost data using approximately common definitions in response to requests from state and federal agencies and legislatures. The underlying causes of the variation in costing procedures are too fundamental to be dealt with merely by seeking common definitions.

Progress undoubtedly can and should be made in the conduct of good comparative cost studies. For some years to come, progress is most likely to be made if the problem is approached as an area for research separated from decisionmaking bearing on the growth, or even survival, of institutions and programs. Any concerted effort to combine multiprogram effectiveness and cost studies is premature until the methodology and procedures for each kind of study are better developed. If such an effort were pursued, the most likely effect will be to impede the necessary prior work in both areas.

Throughout this report we have discussed issues and problems that must be dealt with in the identification of useful information and of

data collection design. We must conclude the report by noting the major dilemma underlying a decision to undertake a study of the scope implied in the original request from NIE.

Good evaluation of educational programs requires thorough longitudinal studies. The time required for such studies exceeds the capability to anticipate decisions required in the political arena of education, however. Investment in a large-scale longitudinal study is questionable when financial resources are limited, if the purpose of the study is to provide information for decisionmakers. Simpler studies using high quality but less detailed data collection devices and designs may serve the purposes of higher education more fruitfully.

REFERENCES

Medsker, L., Edelstein, S., Kreplin, H., Ruyle, J., & Shea, J. *Extending opportunities for a college degree: Practices, problems, and potentials*. Berkeley: Center for Research and Development in Higher Education, University of California, 1975.

Micek, S. S. & Arney, W. R. *The higher education outcome measures identification study*. Boulder, Colo.: National Center for Higher Education Management Systems at WICHE, 1974.

Palola, E. G., Lehman, T., Bradley, A. P., Jr., & Debus, R. *PERC handbook: Program effectiveness and related costs*. Saratoga Springs, N.Y.: Office of Research and Evaluation, Empire State College, 1977.

Smith, S. M., & Smith, M. S. *Pre-design activities for evaluation of postsecondary open-learning programs--2: Identification of policymakers' evaluative information needs*. Washington, D.C.: TEAM Associates, 1975.

APPENDIX

Additional References and Instruments

In addition to the references cited in the report, other publications and instruments or sets of instruments were especially helpful in testing out the information specifications shown in Exhibits A through E. An examination of the instruments would be useful as a base for beginning instrument development for a multiprogram study. The instruments have been used variously in multiprogram and institutional self-evaluation studies. Some have been used by traditional institutions and others by nontraditional institutions or programs; some were designed for cross-sectional studies and others were designed as a series for longitudinal studies. Instruments for use in two-year or four-year colleges are included.

PUBLICATIONS

Bowen, H. R., & Douglass, G. K. *Efficiency in liberal education: A study of comparative instructional costs for different ways of organizing teaching-learning in a liberal arts college.* New York: McGraw-Hill, 1971.

Bower, C. P., & Renkiewicz, N. *A handbook for using the student outcomes questionnaires.* Boulder, Colo.: National Center for Higher Education Management Systems at WICHE, 1974. ED 096 624

Clark, D. G., Huff, R. A., Haight, M. J., & Collard, W. J. *Introduction to the resource requirements prediction model 1.6.* Boulder, Colo.: National Center for Higher Education Management Systems at WICHE, 1973. ED 074 999

Gooler, D. D. Criteria for evaluating the success of nontraditional postsecondary education programs. *Journal of Higher Education*, January/February 1977, *48*(1), 78-95.

Graham, T. M. *Evaluation report, external degree program.* The Consortium of the California State University and Colleges, 1977. (Mimeographed)

Hodgkinson, H. L., Hurst, J., & Levine, H. *Improving and assessing performance: Evaluation in higher education.* Berkeley: Center for Research and Development in Higher Education, University of California, 1975. ED 112 712.

Hodgkinson, H. L., Hurst, J., Levine, H., & Brint, S. *A manual for the evaluation of innovative programs and practices in higher education.* Berkeley: Center for Research and Development in Higher Education, University of California, 1974.

Micek, S. S., Service, A. L., & Lee, Y. S. *Outcome measures and procedures manual, field review edition.* Boulder, Colo.: National Center for Higher Education Management Systems at WICHE, 1975. ED 115142.

Pace, C. R., & Associates. *Higher education measurement and evaluation kit.* Los Angeles: Laboratory for Research on Higher Education, UCLA Graduate School of Education, 1975.

Regional Accrediting Commissions of Postsecondary Education. *Evaluation of Institutions of Postsecondary Education.* Washington, D.C.: Council on Postsecondary Accreditation, 1975.

INSTRUMENTS

Multi-institutional

Bureau of Social Science Research, Inc., Washington, D.C.
EDP Survey of College Graduates, 1977 (56 questions)

California State University and Colleges
Consortium External Degree Program, Student Evaluation (35 questions and four ratings of courses)
Consortium Faculty Questionnaire (19 questions, 2 rating forms)

Center for Research and Development in Higher Education, Berkeley:
Nontraditional Program Study
Student Questionnaire, 1974 (47 questions)
Faculty Questionnaires, 1974 (34 to 45 questions in 12 versions of questionnaires)

College Entrance Examination Board
Student Resource Survey, 1973 (67 questions)

Cooperative Institutional Research Program, UCLA and ACE
- 1976 Student Information Form (38 questions)

National Center for Higher Education Management Systems
- Student Outcomes Questionnaires: Four-Year Colleges and University Series (1977)
  - Entering Student Questionnaire (15 questions)
  - Nonreturning Student Questionnaire (21 questions)
  - Graduating Student Questionnaire (19 questions)
  - Recent Alumni Questionnaire (30 questions)
  - Current Student Questionnaire (in development)
- Student Outcomes Questionnaires: Community Colleges Series (1977)
  - Entering Student Questionnaire (18 questions)
  - Nonreturning Student Questionnaire (20 questions)
  - Graduating Student Questionnaire (19 questions)
  - Former Student Questionnaire (31 questions)

TEX-SIS Follow-up for Community Junior Colleges, Texas Education Agency (1976)
- Student's Educational Intent (6 questions)
- Withdrawal Follow-up
  - From individual courses (4 questions)
  - From college (5 questions)
  - Walk-off (7 questions)
- Nonreturning Student Follow-up
  - Not return for subsequent enrollment (23 questions)
  - Occupational/technical students who do not return (14 questions)
- Graduate Follow-up
  - First-year Graduates (20 questions)
  - Third-year Graduates (19 questions)
  - Fifth-year Graduates (19 questions)
- Employer Follow-up (11 questions)

Institutional Self-Study Instruments

University of Colorado, Boulder
- Freshman Questionnaire (39 questions)
- Student Outcome Questionnaire for Program Completers (25 questions)
- Student Outcomes Questionnaire Supplement (21 questions)
- Alumni Survey (34 questions)

Delaware County Community College
- Confidential Exit Interview (5 questions)
- Questionnaire for Nonreturning Students (19 questions)
- Alumni Questionnaire (50 questions)

Empire State College, PERC
- Student Biographical Inventory (46 questions)
- Student Experience Questionnaire (25 questions)
- Program Completion Questionnaire (38 questions)
- Graduate Followup Questionnaire (41 questions)
- Faculty Questionnaire (30 questions)
- Student Rating Form (self-rating on seven areas of intellectual competence and personal development)
- Faculty Rating Form (rating of student on seven areas as above)

New York Regents External Degree Program
- REDP Participants, 1974-75 Survey (13 questions)
- REDP Alumni Questionnaire, 1974-75 (13 questions)
- REDP Baccalaureate Degree Recipients, 1976 Survey (8 questions)

Development in Progress

American College Testing Program, Aubrey Forrest, Project Director
- College Outcomes Measures Project (COMP)
  - Communicating
  - Solving problems
  - Clarifying values
  - Functioning within social institutions
  - Using science and technology
  - Using art

Educational Testing Service, Jonathan Warren, Project Director
- Academic Competencies in General Education
  - Analytic skill
  - Awareness
  - Synthesizing
  - Written communication